Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA

Cynthia M. Sharma[1], Kai Papenfort[1], Sandy R. Pernitzsch[1], Hans-Joachim Mollenkopf[2], Jay C. D. Hinton[3] & Jörg Vogel[1,#]

[1]Institute for Molecular Infection Biology, Research Centre of Infectious Diseases, University of Würzburg, Germany

[2]Max Planck Institute for Infection Biology, Core Facility Microarray/Genomics, Berlin, Germany

[3]Department of Microbiology, Moyne Institute of Preventive Medicine, School of Genetics & Microbiology, Trinity College, Dublin, Ireland.

Corresponding author:
E-mail: joerg.vogel@uni-wuerzburg.de
Phone: +49-931-3182-576 Fax: -578

Institute for Molecular Infection Biology
University of Würzburg
Josef-Schneider-Str. 2 / Bau D15
97080 Würzburg
Germany

Running title: *GcvB regulon*

This is an Accepted Article that has been peer-reviewed and approved for publication in the *Molecular Microbiology*, but has yet to undergo copy-editing and proof correction. Please cite this article as an "Accepted Article"; doi: 10.1111/j.1365-2958.2011.07751.x

Key words:
GcvB, small RNA, Hfq, post-transcriptional control, translation inhibition, ABC transporter, amino acid metabolism

SUMMARY

GcvB is one of the most highly conserved Hfq-associated small RNAs in Gram-negative bacteria and was previously reported to repress several ABC transporters for amino acids. To determine the full extent of GcvB-mediated regulation in *Salmonella*, we combined a genome-wide experimental approach with biocomputational target prediction. Comparative pulse expression of wild-type versus mutant sRNA variants revealed that GcvB governs a large post-transcriptional regulon, impacting ~1% of all *Salmonella* genes via its conserved G/U-rich domain R1. Complementary predictions of C/A-rich binding sites in mRNAs and *gfp* reporter fusion experiments increased the number of validated GcvB targets to more than twenty, and doubled the number of regulated amino acid transporters. Unlike the previously described targeting via the single R1 domain, GcvB represses the glycine transporter CycA by exceptionally redundant base-pairing. This novel ability of GcvB is focused upon the one target that could feedback-regulate the glycine-responsive synthesis of GcvB. Several newly discovered mRNA targets involved in amino acid metabolism, including global regulator Lrp, question the previous assumption that GcvB simply acts to limit unnecessary amino acid uptake. Rather, GcvB rewires primary transcriptional control circuits and seems to act as a distinct regulatory node in amino acid metabolism.

INTRODUCTION

Amino-acids are the buildings blocks of all proteins and therefore constitute an essential component of all living organisms. In enterobacterial species such as *Escherichia coli* or *Salmonella enterica* serovar Typhimurium (from here on *Salmonella*), amino acids are either synthesized from a broad range of nitrogen sources or imported by specialized transport proteins. The transporters help the cell to use externally available amino-acids directly for protein synthesis (avoiding energy-consuming anabolic pathways) or, following catabolism, as nitrogen and carbon sources. Unsurprisingly, amino acid uptake and metabolism are strictly controlled in enterobacteria, and a plethora of regulatory mechanisms are known to alter gene expression in response to nitrogen availability.

Several DNA-binding proteins respond at the transcriptional level to low levels of amino acids in the environment, including the transcription factor Lrp (Chen *et al.*, 1997) which regulates the expression of at least 10% of all genes in *E. coli* (Tani *et al.*, 2002). In concert with Lrp, the GcvA transcription factor controls the *gcvTHP* operon (Stauffer & Stauffer, 1994); this polycistron encodes the glycine cleavage system that provides C-1 units for the biosynthesis of purines, thymine and methionine, and facilitates most cellular methylation reactions (Kikuchi, 1973). Importantly, GcvA does not only regulate protein-coding genes but also activates *gcvB*, a small noncoding RNA (sRNA) gene which is transcribed divergently from the *gcvA* locus (Urbanowski *et al.*, 2000, Argaman *et al.*, 2001). The *gcvB-gcvA* synteny has been preserved throughout the enterobacterial clade, as well as in more distantly related bacterial species such as *Vibrio* (Sharma *et al.*, 2007, Silveira *et al.,* McArthur *et al.*, 2006), ranking GcvB amongst the top sRNAs with known homologs in Gram-negative bacteria.

GcvB was also shown to belong to the large class of Hfq-associated sRNAs (Zhang *et al.*, 2003, Sittka *et al.*, 2008) whose members generally control *trans*-encoded target mRNAs at the post-transcriptional level. The Hfq protein itself is an RNA chaperone that promotes short and imperfect base-pairing between sRNA and targets; it also maintains the intracellular stability of many sRNAs (Valentin-Hansen *et al.*, 2004). Previous work in *E. coli* and *Salmonella* revealed that GcvB represses multiple target mRNAs, most of which encode amino-acid uptake systems relevant for the utilization of external nitrogen sources. Specifically, GcvB targets the first cistrons of the *dppABCDF* and *oppABCDF* operon mRNAs which encode the major enterobacterial dipeptide and oligopeptide transporters, respectively (Urbanowski *et al*., 2000, Sharma *et al*., 2007, Pulvermacher *et al.*, 2008). Other known targets of GcvB include conserved mRNAs with functions in amino acid transport

(*gltI*, *argT*, *livK*, *livJ*, STM4351, *sstT/ygjU*, and *cycA*), and the *Salmonella*-specific STM4531 mRNA of a putative periplasmic arginine-binding protein (Sharma et al., 2007, Pulvermacher *et al.*, 2009c, Pulvermacher *et al.*, 2009b). Since GcvB is most highly expressed in fast growing cells in rich media (Sharma *et al.*, 2007, Argaman et al., 2001), it was thought that the main function of the RNA was to limit energy-consuming amino acid uptake when nutrients are plentiful.

An intriguing feature of GcvB is the exceptional conservation of two single-stranded nucleotide stretches, termed R1 and R2, which contrasts with otherwise considerable variation in the overall length and secondary structure of predicted GcvB species (Sharma et al., 2007). R1 denotes an array of ~30 single-stranded G and U residues separating the first and second stem-loop in the biochemically-determined solution structure of *Salmonella* GcvB RNA; R2 is the single-stranded decamer ACUUCCUGUA between the third and fourth stem-loop (Fig. 1A). Our previous work showed that the G/U-rich R1 sequence base-pairs to the 5' UTRs of seven ABC transporter mRNAs of *Salmonella* to cause translational repression. The corresponding GcvB binding sites are commonly C/A-rich and might normally act as translational enhancer elements of target genes, as shown for *gltI* mRNA (Sharma et al., 2007). Our discovery that a highly conserved RNA domain, i.e. R1 in GcvB RNA, is employed to recognize multiple, functionally related targets has since been extended to other Hfq-dependent sRNA regulators, for example, OmrA/B or RybB, each of which repress several outer membrane porins through a conserved short target-binding domain (Papenfort & Vogel, 2009, Guillier *et al.*, 2006, Papenfort *et al.*, 2010, Bouvier *et al.*, 2008, Guillier & Gottesman, 2008, Balbontin *et al.*, 2010, Johansen *et al.*, 2006).

Targets of *Salmonella* GcvB were initially identified by proteomic analysis of down-regulated periplasmic proteins and computer-aided predictions of RNA interactions with 5' regions of mRNA (-/+ 50 nt of start codon; Sharma et al., 2007). However, the pleiotropic nature of *E. coli* and *Yersinia gcvB* mutants (McArthur *et al*., 2006, Urbanowski *et al*., 2000) and further biocomputational predictions (Tjaden *et al.*, 2006, Busch *et al.*, 2008) suggested that GcvB might have additional mRNA targets. Comparative mRNA profiling of *E. coli* wild-type versus Δ*gcvB* strains revealed that GcvB affected ~70 genes, of which *sstT* and *cycA* were regulated post-transcriptionally according to results with *lacZ* reporter fusions (Pulvermacher et al., 2009b, Pulvermacher et al., 2009c). However, since neither of the steady-state proteomic or microarray-based transcriptomic analyses distinguished between direct targets and secondary regulatory events of a sRNA, the full target suite of GcvB remained to be identified.

Global mRNA profiling following ectopic sRNA pulse-expression (Masse *et al.*, 2005, Papenfort *et al.*, 2006) predicted targets of several Hfq-dependent regulators with high confidence. This strategy of short sRNA over-expression from a highly inducible promoter such as P_{BAD} effectively circumvents the pleiotropic effects ensuing from constitutive sRNA expression. Thus far, this approach has only been used to pulse-express full-length sRNAs. Here, we describe an extended search for GcvB targets, involving the comparative pulse-expression of GcvB wild-type and mutant sRNAs, complemented by bioinformatics-based predictions of C/A-rich target sites in *Salmonella* mRNAs. This dual strategy identified almost 50 new R1-dependent target candidates, of which thirteen were validated with GFP reporter gene fusions. We show that almost all of these regulations strictly depend on the R1 domain, and also identify a second mechanism involving the exceptionally redundant base-pairing of multiple regions of the sRNA to the *cycA* mRNA. Because CycA is the major glycine transporter, it is possible that its activity might feedback-regulate the biogenesis of GcvB.

RESULTS

Global mRNA profiling upon pulse-expression of GcvB variants

For genome-wide identification of direct GcvB targets, we used a novel comparative pulse-expression approach, assaying the effects of both wild-type and mutant sRNAs. Wild-type GcvB, and variants deleted for the conserved regions R1 or R2 (Fig. 1A), were cloned under control of an arabinose-inducible P_{BAD} promoter, yielding plasmid pBAD-GcvB (pKP1-1), pBAD-GcvBΔR1 (pKP2-6) or pBAD-GcvBΔR2 (pKP30-1), respectively. For confirmation of inducible expression, *Salmonella* wild-type carrying the pBAD control vector (pKP8-35), and *Salmonella* Δ*gcvB* carrying either control vector (Ctr) or the GcvB expression plasmids were grown to mid-exponential phase (OD_{600} of 1) and treated with L-arabinose for up to 15 min (Fig. S1). Northern blot analysis using an oligonucleotide probe directed against the 5' end of GcvB RNA confirmed comparable accumulation of all GcvB variants at the 10 min time-point (Fig. 1B); the plasmid expression yielded six-fold more of GcvB wild-type RNA, and between two- and threefold more of the two mutant RNAs, than the chromosomal *gcvB* gene. For reasons unknown, L-arabinose by itself mildly (~1.5-fold) increased GcvB expression from the chromosome. Note, however, that in the transcriptomic analysis below arabinose was added to both the control and sRNA-inducing samples.

We used whole-genome microarrays to identify relative mRNA expression changes at 10 min of induction in the Δ*gcvB* strain, comparing the pBAD-GcvB, pBAD-GcvBΔR1 or

pBAD-GcvBΔR2 plasmids to the control vector (Table 1). Of the 4,795 *Salmonella* reading frames covered on the microarrays, 54 transcripts were altered >2-fold by at least two sRNA variants (Fig. 1C, Table 1). Consistent with our previous data (Sharma et al., 2007), the known GcvB targets *dppA*, *gltI,* and *oppA* and several co-transcribed cistrons (of the *dppABCDF*, *gltIJKL* and *oppABCDF* operons, respectively) were strongly repressed by the wild-type and ΔR2 forms of GcvB. As expected, these targets were not regulated by the ΔR1 mutant RNA because the deleted R1 region is essential for pairing (Sharma et al., 2007). Weaker repression, again R1-dependent, was observed for other known *Salmonella* targets, namely *livJ*, *argT*, and STM4351, whereas *livK* fell below of the threshold of a 2-fold change expression change (Table S1).

New candidate targets that were repressed in at least two strains were *aroP*, *brnQ, cycA, putP, sfbA, sfbB, yaeC, tppB, yecS, ygjU (sstT)*, and *yifK*, all of which encode amino acid transport proteins, and the *astA, astC, gdhA, lrp, pepN, pheA, serA, trpE,* and *yedO* genes encoding proteins for biosynthesis of individual amino acids. Again, almost all of these regulations were R1-dependent (Table 1; marked in grey). In contrast, the *cycA*, *gdhA* and *lrp* mRNAs were down-regulated by all three GcvB variants, suggesting that their regulation did not strictly require the conserved R1 or R2 regions.

In a reciprocal analysis we looked for loss of regulation by comparing the transcriptome of the *Salmonella* wild-type and Δ*gcvB* strains, after 10 min induction of the pBAD control vector. All previously known GcvB targets but *livK* and STM4351 were up-regulated in the absence of the sRNA (Table S1). Generally, fold-changes were smaller than in sRNA pulse-expression. For example, whilst *dppA*, *gltI* or *oppA* were repressed in a fivefold to 30-fold range by overexpression of GcvB or $GcvB_{\Delta R2}$, the reciprocal up-regulation in Δ*gcvB* versus wild-type strain was only in a 3-fold to 10-fold range. This indicates that, given the threshold of a 2-fold mRNA level change, targets with weak regulation upon GcvB overexpression might not be showing significant up-regulation in the reciprocal sRNA deletion analysis. Nonetheless, the absence of GcvB from the Δ*gcvB* strain generated the expected up-regulation for seven of the new candidate targets identified in the sRNA pulse-expression experiment, namely *astA, astC, cycA, gdhA, ybdH*, *ygjU/sstT*, and *yifK* (Table S1).

Target validation using gfp reporter fusions

Hfq-associated sRNAs such as GcvB generally recognize the 5' region of target mRNAs, and this binding can be assayed in a well-established *gfp* reporter system tailored to monitor post-transcriptional regulation *in vivo* (Urban & Vogel, 2007). We constructed reporter

plasmids for nine new target candidates (Fig. 2A-B), fusing the 5' UTR and up to 15 codons to the N-terminus of GFP. Plasmids pXG-0 and pXG-1 expressing either no *gfp* or the *gfp* gene alone served as the relevant controls (Figs. S2-S4).

Reporters were individually combined in an *E. coli* Δ*gcvB* background with plasmids that constitutively express the wild-type, ΔR1 or ΔR2 variants of *Salmonella* GcvB, or pTP011 as the corresponding control vector (Sharma *et al.*, 2007). The levels and regulations of the various *gfp* fusions were analyzed by three different read-outs: (i) colony fluorescence of bacteria grown on agar plates; (ii) single-cell fluorescence as measured by flow cytometry analysis of overnight cultures; (iii) signal intensity on Western blots probed with an α-GFP antibody to detect the fusion proteins (Figs. 2A-B, S2-S4).

It should be noted that all three assay types have advantages and disadvantages. Analysis of colony fluorescence is the simplest method, but it is not as quantitative as the other two assays. A highly active (fluorescent) reporter gene fusion is needed to reveal regulation by GcvB RNA by both the plate assay and the relatively fast and quantitative flow cytometric approach. In contrast, Western blot analysis can detect regulation even if the fused target protein is non-fluorescent due to, for example, misfolding.

All nine GFP fusions (*ygjU*/*sstT, yaeC, ybdH, gdhA*, *tppB*, *serA, brnQ, lrp,* and *ndk*) showed GcvB-dependent regulation, and the results are summarized in Table 1. In the three assays, the *gfp* alone was not regulated by GcvB, whereas a fusion of the well-established *oppA* target was repressed by both GcvB and GcvBΔR2, and not at all by GcvBΔR1 (Figs. 2A, S2-4). Several new target fusions, namely *ygjU*/*sstT, serA, ndk, ybdH, tppB,* and *yaeC*, showed similar patterns to *oppA* in flow cytometric analysis, which was based on 30,000 single cells per sample (Fig. 2A). In addition, R1-dependent regulation was confirmed on agar plates for *ygjU*/*sstT, yaeC, ybdH, gdhA*, *tppB*, and *ndk* (Fig. S2); Western blot confirmed R1-dependent regulation for *brnQ*, *ygjU*/*sstT, gdhA,* and *serA* (Figs. 2B, S4).

The flow cytometric approach was a particularly valuable tool to validate the identity of GcvB-dependent targets, since the *tppB, yaeC, lrp,* and *ndk* targets were only shown to be regulated on agar plates, and not by western blot. The fluorescence of the *ybdH::gfp* fusion was too low to be faithfully assayed on plates. However, all five fusions showed R1-dependent regulation in the single cell FACS experiments. In addition, using Northern blot analysis, we observed an approximately threefold elevated level of chromosomal *lrp* and *ndk* mRNA levels in a *Salmonella* Δ*gcvB* strain at mid-log phase, as compared to wild-type, which added confidence to our predictions (Fig. S5).

The flow cytometric analysis did not identify R1- or R2-dependent regulation of low-fluorescence fusions including *brnQ::gfp* which for an unknown reason showed little detectable fluorescence, despite the production of stable fusion proteins. However, the Western blot did validate the R1-dependent repression of *brnQ* that had been predicted by transcriptomics (Fig. 2B). These examples show that a combination of assays can be useful for the validation of candidate sRNA targets.

We previously showed that GcvB RNA can directly bind to certain C/A-rich regions in the 5' UTR of ABC transporter mRNAs. Similar C/A-rich regions for base-pairing with the G/U-rich R1 element of GcvB can be predicted in the 5' UTRs of the new targets *ygjU*/*sstT*, *yaeC, gdhA, ybdH, serA, ndk,* and *brnQ* (Fig. 2C). In addition, *tppB,* for which *gfp* fusion regulation was visible on agar plates and by flow cytometric analysis, contains two potential C/A-rich target sites, one located in the early CDS (coding sequence) and the other ~50 nucleotides upstream of the start codon.

Biocomputational predictions reveal additional GcvB targets

The microarray-based approach cannot identify GcvB targets that are either not expressed under the tested growth condition or controlled at the level of translation without dramatic changes at the mRNA level; for example, the known *livK* target (Sharma *et al.*, 2007) showed only a 1.5-fold repression, below the threshold for significant regulation (Tables 1 & S1). Computer-aided searches for sRNA complementarity can predict mRNA targets in an expression-independent manner, but typically suffer from a high number of false positives especially with sRNAs such as GcvB which have a high degree of redundancy within their recognition sequences. Although most targets required the R1 domain, we searched the *Salmonella* genome for additional GcvB targets with potential C/A-rich binding sites.

We first defined a consensus motif for the previously-validated GcvB target sites, building upon the previously reported *argT*, *dppA, gltI*, *livJ, livK, oppA,* and STM4351 target sites (Sharma et al., 2007) and the nine putative R1 recognition sites in the *brnQ, gdhA, lrp, serA*, *ndk, ybdH, yaeC, tppB,* and *ygjU*/*sstT* targets validated above. We used the MEME motif identification software (Bailey & Gribskov, 1998) with parameters set such that up to ten motifs with a length between 6 and 25 nucleotides would be predicted within the cloned fusion fragments as input sequences; the size limit was based on the overall length of R1 which is 24 nucleotides (Sharma et al., 2007). The best motif identified by MEME was 8 nucleotides long, present in all input sequences (Fig. 3A-B), and overlapped with almost all of the previously proposed GcvB interaction sites (Sharma et al., 2007 and Fig. 2C).

Exceptions were *lrp,* for which no significant interaction could be predicted, and *tppB,* where the motif overlaps with the second proposed interaction that is located in the 5'UTR (Fig. 2C). Thus, despite being unrelated in overall mRNA sequence, the known GcvB targets share a common motif with substantial complementarity to R1.

Next, we exploited the above C/A-rich motif in weight matrix-based MAST searches (Bailey & Gribskov, 1998) to predict GcvB binding in the -70 to +30 region (relative to start codon) of all 4,424 annotated *Salmonella* ORFs (McClelland *et al.*, 2001). Sequences matching to the motif with an E-value ≤200 were found in 234 of the 4,424 input mRNAs (Table S2), including 14 of the 16 GcvB sites originally used to define the motif (the motif sites within *lrp* and *gdhA* showed E-values above the cut-off). Next, the *TargetRNA* software (Tjaden, 2008) was used to predict base pairing based on free-energy minimization between the above -70/+30 regions and an extended GcvB pairing region comprising the 24 nucleotide R1 sequence with 12 upstream and 13 downstream flanking nucleotides. A p-value cut-off of 0.05 revealed a list of 42 targets including the validated *dppA, argT, oppA, serA, ybdH,* STM4351, and *yaeC* mRNAs (Table S3); remarkably, a third of them (14 targets) are involved in amino acid transport and metabolism. A C/A-rich motif with an E-value <200 (Tables S2 & S3) is found in 19 of the *TargetRNA* predictions, and half of these in validated targets. Intriguingly, *TargetRNA* again predicted the *yifK* and *trpE* genes as targets; GcvB-dependent regulation of these two mRNAs had been detected on microarrays, yet could not be confirmed with the *gfp* fusions due to low reporter fluorescence.

We selected five of the new amino acid-related target candidates—*mltC*, *ilvC*, *thrL*, *iciA*, and *ilvE*—for validation with *gfp* fusions. The *mltC::gfp* fusion had insufficient reporter activity, but the other four targets were fully validated using the three types of measurement described above. All targets were clearly down-regulated on agar plates (Fig. 2S), and confirmed to have the expected R1-dependence by flow cytometry (Fig. 3C). Regulation of *ilvC, thrL* and *iciA* was also visible on Western blots (Fig. S4). The predicted R1 sites in *iciA*, *ilvC*, *ilvE* and *thrL* were consistently located in the 5' UTR (Fig. 3C), at positions where GcvB is well-established to control mRNA expression (Sharma *et al*., 2007). The *ilvC* and *ilvE* genes function in branched-chain amino acid metabolism or biosynthesis, respectively; *iciA* encodes the transcriptional activator of the arginine transport system; and *thrL* belongs to the threonine (*thr*) biosynthesis operon. The fact that none of these targets showed altered mRNA levels in our transcriptomic experiments (Table 1 & S1), suggests that comprehensive identification of sRNA targets requires a combination of an experimental and a computational approach.

GcvB targets the thrL *leader peptide of the threonine biosynthesis operon*

The newly identified *thrL* target was particularly interesting as it is a regulatory RNA element itself. Specifically, *thrL* encodes a 21 amino-acid leader peptide of the threonine biosynthetic operon (Fig. 4A), translation of which is sensitive to cellular threonine and isoleucine levels and regulates the downstream cistrons by an attenuation mechanism in *E. coli* (Lynn *et al.*, 1982): low levels of the corresponding charged tRNAs stall ribosomes at the eighth threonine or fourth isoleucine codon, preventing the formation of a termination structure and allowing RNA polymerase to continue transcription to the downstream *thrABC* biosynthesis genes.

The predicted C/A-rich R1 binding site of GcvB site is located immediately upstream of the *thrL* start codon (Figs. 3C and 4A). However, owing to the numerous threonine codons, the leader peptide region itself is C/A-rich. To test whether both regions are recognized by GcvB, we tested *gfp* fusions to either the 19th or 5th codon of *thrL*, which do or do not include the threonine codons, respectively. Both fusions were comparably regulated by GcvB in an R1-dependent manner, suggesting that only the upstream site is targeted (Fig. 4B). Moreover, a deletion of the upstream C/A-rich site rendered the longer fusion refractory to GcvB. These genetic results not only confirmed the biocomputational target site prediction for *thrL*, but also argue that only certain C/A-rich sequences can serve as a GcvB target site.

GcvB represses translation of the glycine transporter CycA

In our transcriptomic experiments, the *cycA* mRNA was ~5-fold-repressed by all three GcvB variants, suggesting that the regulation of *cycA* did not rely solely on R1 or R2 and involved alternative binding sites (Table 1, Fig. 1C). The fact that *cycA* mRNA levels were 3.2-fold increased in the *Salmonella* Δ*gcvB* versus wild-type strain (Table S1 and quantitative RT-PCR data not shown), and that CycA acts as a permease for glycine, D-alanine, D-serine, and D-cycloserine in *E. coli* (Cosloy, 1973, Robbins & Oxender, 1973, Russell, 1972, Wargel *et al.*, 1971), increased confidence in *cycA* being a direct GcvB target.

We confirmed both the repression of *cycA* and its apparent independence of R1 and R2 with a *cycA::gfp* fusion (Fig. 5A). To narrow down potential interaction regions in the sRNA, we constructed four additional GcvB mutants and tested their effects on *cycA::gfp* by both flow cytometric analysis and the colony fluorescence assay (Fig. S7). Strikingly, neither a R1/2 double deletion, nor 5' or 3' truncations abrogated the ability of GcvB to repress the GFP reporter, suggesting that GcvB possess several redundant pairing regions for *cycA*. To

rule out the possibility that the multi-copy GcvB DNA facilitates the repressive effect, for example by antagonizing a protein required for *cycA* mRNA expression, we deleted the promoter from GcvB expression plasmid, pP_L*gcvB*. This finally abrogated the repression of *cycA::gfp*, arguing that *cycA* is indeed regulated by GcvB RNA and not via the DNA of the *gcvB* gene.

The *cycA* mRNA contains a predicted C/A-rich motif with a conserved location near its ribosome binding site (Fig. 5B, Table S2). To test whether antisense pairing of the R1 domain of GcvB with this motif would prevent ribosome loading, 30S ribosome toeprinting assays were performed (Hartz *et al.*, 1988) using a 5' mRNA fragment of a *cycA* reporter gene (*gfp* fused to codon 10 of *cycA*). A 5' end-labeled primer was annealed to the *gfp* coding region, and incubated with 30S subunits in the presence or absence of uncharged $tRNA^{fMet}$, followed by cDNA synthesis. Analysis of the extension products (Fig. 5C, lanes 1-3) revealed one ribosome-induced 'toeprint'—a $tRNA^{fMet}$-dependent termination site—at the expected +15/+16 position (start codon A is +1), proving that the start codon was correctly annotated.

This toeprint signal was decreased in the presence of increasing concentrations of wild-type GcvB RNA (lanes 4-6), suggesting antisense-specific inhibition of 30S binding. By contrast, the unrelated MicA RNA (lane 10) did not inhibit ribosome association. Strikingly, inhibition was again observed with ΔR1 and ΔR2 variants of GcvB RNA, and a 3' truncation mutant (3'Δ) which also lacks R2 (lanes 7-9). Specifically, a 10-fold excess of GcvB wild-type reduced the binding of the 30S subunit to *cycA* by approximately sevenfold, whereas the mutant RNAs were observed to decrease the signal between threefold and fivefold; we assume that some of these variations result from different binding affinities of the tested GcvB variants to *cycA* mRNA. Nevertheless, the *in vitro* experiment argues that several regions in GcvB can independently inhibit translation initiation of the *cycA* target.

Structural mapping of GcvB-cycA RNA interactions

To map potential GcvB interaction sites, an *in vitro*-synthesized fragment of the *cycA* leader was subjected to RNA structure probing with RNase T1 or lead(II) acetate, either alone or in the presence of GcvB wild-type or mutant RNAs. Protections were most visible for cleavage by lead(II) acetate (Fig. 6A). The full-length GcvB, ΔR2 and 3'Δ mutant RNAs protected ~7 nucleotides of the C/A-rich motif (Fig. 6A, indicated by blue vertical lines). In line with our hypothesis that this motif is targeted by R1, protection was not seen with ΔR1, the R1/R2 double deletion mutant (ΔR1ΔR2), or an additional 5' truncation mutant (5'Δ). The GcvB

WT, ΔR1, 5'Δ and 3'Δ RNAs also protected the RNA around the Shine-Dalgarno sequence (SD) of *cycA*; with ΔR1 and the 5' truncation this protection extends to ~18 nucleotides starting from position -7 relative to AUG (Fig. 6A, red bars). The protection was less prominent for ΔR2 and ΔR1ΔR2 mutants.

In a complementary experiment with labeled *cycA*, we subjected the six GcvB-*cycA* complexes to the double strand-specific RNase III (Fig. 6A, right panel). Wild-type GcvB, ΔR1, and 5'Δ induced strong R1-independent cleavage in the *cycA* SD around position -8, whereas the cleavage was lost for the ΔR2, the ΔR1ΔR2, and the 3' truncation mutants. This cleavage indicates duplex formation, and together with the probing results described above we tentatively predict that the 5' region (involving R1) and 3' region (involving R2) of GcvB can independently pair with the 5' UTR of *cycA* (Fig. 6C). All three tested GcvB variants promote further RNase III cleavage in the upstream 5' UTR of *cycA* (position -44) that might involve the stem-loop 2 of GcvB which was contained in all six tested RNAs.

The reciprocal probing of labeled GcvB RNA in the presence of the *cycA* RNA fragment indicated multiple *cycA* binding sites on GcvB RNA which were supported by strong RNase III cleavages (Fig. 6B). Importantly, these experiments favor our model of independent interactions of the extended R1 or R2 regions of GcvB (interactions marked by blue or red vertical bar, respectively, in Fig. 6C). The first interaction involves the C/A-rich site upstream of the SD, whilst the other overlaps the SD and start codon of *cycA*.

DISCUSSION

Early research on bacterial sRNAs often focused on an individual target mRNA; for example, *ompF* mRNA has remained the only investigated target of the prototypical MicF sRNA in 25 years (Mizuno *et al.*, 1984). Since then, many studies in *E. coli*, *Salmonella* and *Staphylococcus aureus* have revised the perception of sRNAs as dedicated regulators of a single gene. It has now become clear that sRNAs can entirely rewire primary transcriptional circuits by acting on multiple *trans*-encoded targets (Papenfort & Vogel, 2009, Gottesman & Storz, 2011). This finding explains the results of global analyses of Hfq-bound transcripts, which identified many more potential mRNA targets than sRNA regulators (Zhang et al., 2003, Sittka et al., 2008, Sittka *et al.*, 2009). However, the identities of these emerging post-transcriptional regulons remained to be defined.

GcvB was one of the first sRNAs shown to mediate direct antisense control of a large number of functionally related mRNAs whose encoded proteins belonged to a single mechanistic class, *i.e.* the ABC transporters of amino acids (Urbanowski et al., 2000,

Pulvermacher et al., 2009b, Pulvermacher et al., 2009c, Sharma et al., 2007). A similar enrichment, though at the level of cellular function only, is visible for RyhB sRNA of *E. coli* which controls the synthesis of diverse proteins involved in iron homeostasis (Masse et al., 2005, Masse & Gottesman, 2002, Vecerek *et al.*, 2007, Prevost *et al.*, 2007, Desnoyers *et al.*, 2009, Salvail *et al.*, 2010). Likewise, functional or mechanistic relationships in target suites have been observed for MicA, OmrA/B, and RybB (repression of outer membrane protein synthesis; (Vogel & Papenfort, 2006, Guillier et al., 2006, Guillier & Gottesman, 2008, Papenfort et al., 2010), FnrS and Spot42 (metabolic enzymes; (Durand & Storz, Bohn *et al.,* Chevalier *et al.*, 2009, Beisel & Storz, 2011, Boysen *et al.*), as well as RNAIII (virulence factors (Huntzinger *et al.*, 2005, Boisset *et al.*, 2007).

Regarding ABC transporter regulation, RydC of *E. coli* and AbcR1 of *Agrobacterium tumefaciens* were reported to specialize in this function (Antal *et al.*, 2005, Wilms *et al.*, 2011). It will be interesting to see how additional RNAs connect to the GcvB network and whether sRNAs from more distant bacteria have evolved similar regulatory networks to the GcvB RNA. The most similar case reported so far is RsaE, a conserved sRNA in *S. aureus* and other Gram-positive species which has no apparent homology to GcvB, and targets both *oppA* and metabolic genes (Bohn et al., 2010, Chevalier et al., 2009).

The present study used a new combination of experimental pulse-expression and biocomputational motif hunting to determine the post-transcriptional regulon of GcvB in *Salmonella*, and to identify the role of the highly conserved R1 and R2 domains in facilitating the global activity of this sRNA. Our transcriptomic analysis showed that a short pulse of GcvB expression significantly altered the mRNA expression levels of 54 genes (Table 1), corresponding to ~1% of all protein-coding genes of *Salmonella*. Correction for likely co-regulation of polycistronic genes such those of the *dpp*, *glt* or *opp* operons reduces the list to 45 GcvB-regulated cistrons. The present count of independently validated *Salmonella* GcvB targets runs at 21 mRNAs, which is currently—to the best of our knowledge—the largest number of validated targets for a single sRNA. Although our target confirmation had a certain functional bias, it is clear that GcvB plays a major role in the control of amino acid uptake and metabolism of *Salmonella* (Fig. 7). Of the new targets, *cycA* is particularly interesting as it shows conserved regulation in *E. coli* and encodes the major glycine transporter (Robbins & Oxender, 1973, Ghrist & Stauffer, 1995, Pulvermacher et al., 2009b). The Stauffer lab showed that transcriptional regulation of the *E. coli gcvB* promoter is itself glycine-dependent; high levels of glycine in the medium promote GcvA-mediated activation, whereas low levels favour GcvA/GcvR-mediated repression (Stauffer & Stauffer,

2005). Thus, the repression of glycine import via CycA could facilitate a negative feedback loop that controls the synthesis of regulator GcvB.

Importantly, the nature of the new *Salmonella* targets, combined with recent results in *E. coli* (Pulvermacher et al., 2009b, Pulvermacher et al., 2009c), challenge the previous assumption that GcvB simply serves to limit unnecessary amino acid uptake under nutrient-rich conditions (Sharma et al., 2007, Pulvermacher et al., 2009c). For branched chain amino acids (leucine, isoleucine, and valine), GcvB now appears to concomitantly repress genes of both ABC transporters for import (*brnQ*, *livJ*, and *livK*) and the enzymes for *de novo* biosynthesis (*ilvC* and *ilvE*). A similar pattern is observed for arginine, glutamate and serine (import/biosynthesis: *argT*/*iciA*, *gltI*/*gdhA*, and *ygjU*(*sstT)*/*serA,* respectively). Thus, GcvB shuts down both the import and biosynthetic machinery that provide bacterial cells with amino acids.

The concomitant post-transcriptional repression of Lrp, a transcription factor with dual function, would amplify the proposed shut-down function of GcvB: Lrp is an activator of many genes that function during famine, and represses genes that function during a feast (Calvo & Matthews, 1994). Expression of Lrp is induced by the alarmone ppGpp (Landgraf *et al.*, 1996) during transition into stationary phase, and Lrp regulates the expression of at least 10% of all genes in *E. coli* (Tani *et al.*, 2002). The combination of these repressive functions may explain why high expression of GcvB is detrimental for *Salmonella* (Sharma et al., 2007). More work is clearly required to understand the precise physiological role of GcvB; because the *gcvB* gene is physically linked to *gcvA* in many species beyond *E. coli* and *Salmonella* (Sharma et al., 2007), its putative metabolic function is likely to be widely conserved. In addition, we note a strong overlap of repressed target genes in *Salmonella* with those observed in *E. coli* K12 (Pulvermacher et al., 2009b). Although 25% of the protein-coding genes differ between *Salmonella* and *E. coli* K12 (Porwollik & McClelland, 2003), most of the regulated genes are widely conserved (Table 1). Unlike other highly conserved Hfq-associated sRNAs which readily integrated *Salmonella*-specific mRNAs into their smaller target suites (Papenfort et al., 2010, Papenfort *et al.*, 2009), GcvB has rarely done so: of the 45 GcvB-regulated genes, only STM4351 and STnc310 (Table 1) are specific to *Salmonella* bacteria, whereas the other STM genes are absent from *E. coli* K12, but present in other enteric bacteria. These comparative genomic findings support our suggestion that GcvB specialises in a process that is fundamental to bacterial physiology.

In stark contrast with the *E. coli* situation, we found no *Salmonella* genes that were significantly activated by pulse-expression of GcvB, or conversely, that showed reduced

expression in *Salmonella* Δ*gcvB;* by contrast, the Stauffer lab identified many down-regulated genes in *E. coli* Δ*gcvB* (Pulvermacher et al., 2009b). Similarly, we did not observe a previously reported up-regulation of *rpoS* mRNA encoding a major stress sigma factor which had been postulated to underlie a GcvB-dependent ability of *E. coli* to survive low pH (Jin *et al.*, 2009). At this point we can offer no explanation for these species-specific differences, but conclude from our own results that GcvB might exclusively act as repressor of mRNA. This is remarkable since other sRNAs that repress many targets, for example, ArcZ, DsrA, RNAIII or RyhB, also activate at least one target mRNA (Mandin & Gottesman, Majdalani *et al.*, 1998, Lease *et al.*, 1998, Novick *et al.*, 1993, Prevost et al., 2007).

Our study proves that the R1 domain of GcvB is the major determinant of mRNA regulation. Deletion of R1 abrogated almost all mRNA repression in the pulse-expression experiment (Fig. 1C), and most of remaining targets (*gdhA*) showed R1-dependence when tested independently with *gfp* reporter fusions (Figs. 2, S2-S4). The crucial role of R1 reflects the emerging importance of conserved short domains for multiple-target regulation. Recent studies by us and others have identified conceptually similar domains in numerous sRNAs. For example, a 16 nucleotide stretch in RybB is sufficient to down-regulate many outer membrane proteins under conditions of envelope stress (Balbontin *et al.*, 2010, Papenfort et al., 2010), and a 10 nucleotide stretch in SgrS base-pairs with the sugar transport *ptsG* or *manX* mRNAs (Vanderpool & Gottesman, 2004, Rice & Vanderpool, 2011, Kawamoto *et al.*, 2006). R1 stands out amongst these domains because it has the highest risk for promiscuous base-pairing; each position in this string of almost thirty G and U residues permits two different base-pairs (G:C or G:U, and U:A or U:G, respectively). Theoretically, this redundancy could facilitate millions of different GcvB-mRNA contacts, especially when the number of mismatches observed in bona fide GcvB pairing is considered (Fig. 2C). However, base-pairing by GcvB is highly selective and requires a CACAaCAY core motif in the target (Fig. 3B). The identification of the core motif greatly improved target predictions and revealed additional R1-dependent GcvB-regulated genes whose mRNAs were either too poorly expressed or insufficiently destabilized by GcvB for detection by sRNA pulse-expression. Whether the motif always functions independently as a translational enhancer element, as previously shown for *gltI* mRNA (Sharma et al., 2007), remains an intriguing question.

In more general terms, the abrogated mRNA repression of the ΔR1 mutant of GcvB suggests that the inclusion of sRNA variants with mutations in conserved regions could significantly improve experimental target discovery. This would not necessarily require the

deletion of an entire domain (which bears the risk of sRNA destabilization); for example, a point mutation in the conserved R16 domain of RybB sRNA abrogates the regulation of the majority of known targets without impacting on sRNA biogenesis and stability (Papenfort et al., 2010). However, we also discovered a limitation of the combined biocomputational and transcriptomic approach because no strictly R2-dependent targets were detected (Fig. 1C), despite the fact that R2 is the most highly conserved region in GcvB (Sharma et al., 2007). We speculate that R2 mediates regulation at the translational level without impacting upon mRNA levels. Recent developments in LC-MS-MS technology would permit the determination of global changes in protein synthesis in response to sRNA pulse-expression.

Besides the strictly R1-dependent targets, both our microarray approach and the GFP reporter gene assays indicated that there are some targets, e.g. *cycA* and *lrp* mRNA, which are still regulated by the GcvBΔR1 mutant, predicting that these regulatory interactions involve base-pairing by multiple redundant regions of GcvB RNA. We have previously used an *in vitro* translation assay (reconstituted 70S ribosomes), Hfq and *in vitro* transcribed GcvB and target RNAs to demonstrate that the *gltI* mRNA is repressed in a strictly R1-dependent manner. Our preliminary results using the same system with the *cycA* and *lrp* mRNAs show that both of these targets are repressed by GcvB regardless of the singular or combined absence of R1 and R2 (Figs. 8 and S6). None of the GcvB variants impacted translation of the control *ompD* mRNA, which argues that the observed effects are specific. Although this *in vitro* translation experiment does not reveal which regions of GcvB regulate *cycA* and *lrp in vivo*, it does provide strong evidence that GcvB regulates these two targets at the level of translation.

What may such multiple interactions with R1-independent target RNAs look like? For example, GcvB pairing to the *cycA* mRNA is likely to involve the R2 domain, at least according to our *in vitro* probing results: there are contacts of R2 with the extended SD region of *cycA* (Fig. 6), and GcvB RNAs devoid of R2 (ΔR2, 3'Δ) less efficiently repressed 30S binding of this target (Fig. 5C). However, these putative contributions of R2 to *cycA* regulation were not borne out *in vivo,* and seem to be diminished by exceptionally redundant base pairing through other GcvB regions. In fact, GcvB wild-type, ΔR1 and ΔR2 showed similar levels of down-regulation of *cycA* in both the pulse-expression and *gfp* fusion experiments (Figs. 1C and 5A), and even an R1/R2 double-deletion mutant was fully capable of regulating *cycA::gfp* (Fig. S7). As such, our results are consistent with the absence of a single essential target site, as determined by mutation of a chromosomal *cycA::lacZ* fusion of *E. coli* (Pulvermacher et al., 2009b). Interestingly, *E. coli* GcvB was recently shown

to interact with *oppA* mRNA via both the consensus R1 and a second interaction site; however, this second near-perfect 10-bp duplex did not affect *oppA* regulation *in vivo* (Busi *et al.*, 2010). Altogether, *cycA* regulation *in vivo* likely involves multiple regions of GcvB.

Although an increasing number of Hfq-associated sRNAs use a single conserved domain for multiple targets (Masse & Gottesman, 2002, Holmqvist *et al.,* Guillier & Gottesman, 2008, Pfeiffer *et al.*, 2009, Papenfort et al., Balbontin et al., Lease et al., 1998, Majdalani et al., 1998), double-contacts have been known in the pairing of *E. coli* OxyS with *fhlA* mRNA (Argaman & Altuvia, 2000) and the target interactions of *Staphylococcus* sRNAs (Chabelskaya *et al.*, 2010, Boisset et al., 2007). Moreover, *E. coli* FnrS sRNA was recently discovered to use two separate single-stranded regions for different sets of targets (Durand & Storz, 2010), yet, unlike with GcvB and its *cycA* and *lrp* targets, both of the two domains of FnrS are essential for each associated target set.

We have identified another exciting GcvB target, the *thrL* leader peptide region upstream of *thrABC*. This target was not captured in our transcriptomic experiments but identified biocomputationally via its C/A-rich binding site and subsequently validated by *gfp* fusions (Figs. 3 & 4). Re-inspection of the transcriptomic data for <2-fold regulation, however, suggests that by targeting *thrL*, GcvB might also mediate low-level repression of the downstream *thrABC* genes (≥1.2-fold), at least at the level of mRNA. This co-repression of *thrABC* with *thrL* resembles the recently described regulation of *fur* mRNA in *E. coli* where RyhB sRNA targets an upstream ORF (*uof*) to which the translation of *fur* is strictly coupled (Vecerek et al., 2007). Unlike the *uof* of *fur*, however, *thrL* is itself a regulatory RNA, and belongs to the class of attenuators thought to act autonomously to control *cis*-encoded downstream genes. Our study identifies the first case of an sRNA that actually targets an attenuator region, suggesting that the complexity of RNA-mediated control in *cis* can be enhanced by the action of a *trans*-encoded sRNA. The involvement of Hfq/sRNA-mediated regulation is supported by our observation that *thrL* is co-immunoprecipitated with Hfq protein in log-phase *Salmonella* (Y. Chao & J. Vogel, unpublished results). Further inspection of Hfq co-immunoprecipitation datasets for *Salmonella* and *E. coli* suggested that additional attenuator regions, for example *ivbL* (the *ilvB* operon leader peptide), might be controlled by sRNAs in a similar fashion (Zhang et al., 2003, Sittka et al., 2008, Sittka et al., 2009).

Has the combinatorial strategy of pulse-expression and biocomputational motif search saturated the target search for GcvB in *Salmonella*? We consider this unlikely given that some mRNAs such as *yifK* were both clearly down-regulated by pulse-expression (Fig. 1C) and predicted to possess a reasonable interaction region (Table S2), but could not be

validated with a *gfp* fusion due to low reporter activity (unpublished results). Moreover, some targets that were tested with GFP fusions showed inconsistent regulation with the three different readout methods (colony fluorescence on agar plates, Western blot and flow cytometric analysis). Our two-plasmid system relies upon ectopic over-expression of target fragments, and it is possible that successful regulation by GcvB might only be seen when the full-length mRNA accumulates to a lower physiological concentration. The target fusions could lack recognition sites for Hfq and RNase E, or structural elements that enable efficient GcvB binding as were found in OxyS-*fhlA* (Argaman & Altuvia, 2000). In this study, the 5' regions of target candidates were cloned according to our mapping of transcriptional start sites (5' RACE) or promoters derived from literature. This does not exclude the possibility of alternative transcriptional start sites yielding longer or shorter mRNAs which are the genuine targets of GcvB. We considered whether GcvB might target the CDS, as was recently reported for other Hfq-dependent sRNAs (Bouvier et al., 2008, Papenfort et al., 2010, Pfeiffer et al., 2009); the currently mapped interactions show that GcvB acts almost exclusively in the 5' UTR, at both the RBS and upstream (Sharma et al., 2007; Figs. 2-3).

Based on the following criteria: 1) observed regulation in our combined microarray analyses, 2) validation of post-transcriptional regulation in all reporter gene assays, and 3) a predicted C/A-rich target site, we believe that at least *sstT*, *ybdH*, *serA*, *gdhA*, and *cycA* are direct GcvB targets. In the case of *brnQ*, GFP reporter gene regulation was only observed on Western blots due to low fluorescence of the fusion. Since it meets all the other criteria, we also consider *brnQ* to be a direct GcvB target. In addition, the presence of an interaction with a C/A-rich site and the regulation of the GFP reporter gene fusions of the predicted targets, *ilvC*, *thrL*, *iciA*, and *ilvE*, supports these four mRNAs as direct targets. Furthermore, our *in vitro* translation experiments show that *cycA* and *lrp* are direct targets (Fig. 8). Two targets, *ndk* and *yaeC,* meet all of the above mentioned criteria, except for regulation on Western blots. Overall, we consider that the flow cytometric analysis is superior to Western blots, as it assays fluorescence at the level of single bacterial cells. Combined with the high confidence prediction of interaction sites, we also regard *ndk* and *yaeC* as direct targets. The only problematic target is *tppB*: although its mRNA was regulated in transcriptomic experiments, this was not consistently validated by reporter gene fusions and no convincing single target site was predicted within the *tppB* 5'UTR.

The overall count of new bona fide targets for GcvB is 13 (out of 14 candidates). One explanation for possible indirect effects might be the pleiotropic impact of constitutive over-expression of sRNAs for example, through titration of Hfq binding (Papenfort et al.,

2009, Hussein & Lim, 2011, Zhang *et al.*, 1998). In one scenario, Hfq perhaps in concert with an unrelated sRNA could stabilize an mRNA, and the titration of Hfq by GcvB would result in essentially the same down-regulation as proposed for direct regulation by GcvB. However, we note that many GcvB targets are up-regulated in Δ*hfq* strains of *Salmonella* and *E. coli* (*dppA*, *oppA*, *gltI*, STM4351, *ilvC*, *gdhA*, *serA*, *lrp*, *yaeC*, *ybdH*, *sstT*, and *ndk* (Sittka *et al.*, 2007, Guisbert *et al.*, 2007, Ansong *et al.*, 2009, Sittka et al., 2008), supporting our model that GcvB generally acts as a repressor via Hfq.

In summary, given that deep sequencing analysis of Hfq-associated *Salmonella* transcripts identified an excess of potential targets over putative regulator RNAs, i.e. >700 mRNAs versus ~100 sRNAs (Sittka et al., 2008), some sRNAs must have many targets and control large regulons. The data accumulated thus far strongly argue that GcvB is one such Hfq-dependent riboregulator in *Salmonella which can* impact on ~1% of all mRNAs. Therefore, the research into this sRNA holds promise to improve our understanding of how riboregulation achieves a similar global outcome as do transcription factors.

ACKNOWLEDGEMENTS

We thank Karsten Hokamp for designing the SALSIFY microarray, and Franziska Seifert and Ina Wagner for technical assistance. We are grateful to Sacha Lucchini for help at an early stage of this project. J.C.D.H. was supported by a Principal Investigator award (reference 08/IN.1/B2104) from Science Foundation Ireland. The Vogel lab was supported by the DFG (German Research Council) Priority Program SPP1258 *Sensory and Regulatory RNAs in Prokaryotes* SPP1258, Vo875/3-1, and BMBF (German Ministry of Education & Research) grant 01GS0806/JV-BMBF-01.

FIGURE LEGENDS

Figure 1: Pulse-expression of GcvB wild-type and mutant RNAs identifies the roles of the R1 and R2 domains.

A) Schematic drawing of the secondary structure of GcvB wild-type RNA. Dotted lines in the bars below the structure denote regions that were deleted in the GcvB mutants. The highly conserved regions, R1 and R2, which are deleted in the mutant RNAs are shown in red in the simplified GcvB structure on the left.

B) pBAD-inducible expression of GcvB variants**.** GcvB wild-type (GcvB) and GcvBΔR1 (ΔR1, deletion of residues 66-89) or GcvBΔR2 (R2, deletion of residues 136-144) mutant RNAs were expressed in *Salmonella* SL1344 Δ*gcvB* from the P_{BAD}-inducible promoter by addition of L-arabinose (0.2% final concentration) to cultures grown to an OD_{600} of 1.0. Samples were taken prior to (0 min) and 10 min after induction. Northern Blot analysis indicates strong induction of GcvB wild-type and mutant RNAs after 10 minutes of induction. Samples taken from *Salmonella* wild-type harbouring a pBAD control plasmid (Ctr) revealed that chromosomal GcvB is ~1.5-fold increased after 10 min of arabinose addition. Duplicate samples of the *Salmonella gcvB* mutant carrying either the pBAD control (Ctr), pBAD-GcvB wild-type (GcvB), pBAD-GcvBΔR1 (ΔR1), or pBAD-GcvBΔR2 (ΔR2) plasmid were grown to an OD_{600} of 1.0, and RNA samples were taken after 10 min of L-arabinose induction (0.2% final concentration).

C) The expression profiles of genes regulated by pBAD-GcvB, -GcvBΔR1 or -GcvBΔR2 when compared to pBAD control plasmid. Genes that displayed differential expression (>2-fold) in at least two conditions are shown. Each horizontal bar represents the expression level of a single gene. Green shows that the gene is down-regulated, black that it is expressed at similar levels and red that it is up-regulated.

Figure 2: Proposed duplexes of new GcvB target interactions and validation with GFP reporter gene fusions.

A) *E. coli* Δ*gcvB recA*⁻ strains carrying the *oppA::gfp*, *ygjU(sstT)::gfp*, *serA::gfp*, *yaeC::gfp*, *ndk::gfp*, *gdhA::gfp*, *tppB::gfp*, *ybdH::gfp*, or *lrp::gfp* fusion plasmids in combination with control vector pTP11 (black), or plasmids expressing *Salmonella* wild-type GcvB RNA (pP_L*gcvB*, red) or two of the mutant alleles (pP_L*gcvB*$_{\Delta R1}$, blue; pPL*gcvB*$_{\Delta R2}$, green) were grown to stationary phase and were subjected to flow cytometry analysis. All data acquired from the experiments are plotted in fluorescence histograms generated from all events measured (30,000 events). Cellular fluorescence is given in arbitrary units (GFP intensity). Regulation by GcvB wild-type or GcvBΔ_{R2} is visible as a shift of the fluorescence curves to the left, towards lower GFP intensities.

B) Western blot of GFP alone and BrnQ::GFP fusion protein prepared after growth for 14 hours to stationary phase from *E. coli* Δ*gcvB recA*⁻ carrying the indicated plasmids as in A. GroEL was probed as loading control. Fold changes of BrnQ::GFP fusion protein levels (upon normalization to GroEL levels) by *gcvB*, *gcvB*Δ_{R1}, or *gcvB*Δ_{R2} co-expression relative to the control plasmid were: -2.0/ -1.1/ -2.0.

C) Proposed RNA duplexes formed by GcvB with eight new target mRNAs: *ygjU(sstT), yaeC, gdhA, ybdH, serA, ndk, brnQ* and *tppB*. Positions in the target sequences are given as distance to the mRNA start codon. C/A-rich motifs are highlighted in blue in the interactions. For *tppB* two interactions are predicted, with the first one being located in the coding sequence.

Figure 3: Definition of a C/A-rich motif and bioinformatic-based prediction of additional GcvB targets

A) Location of the predicted C/A-rich motif in the input sequences.

B) A consensus motif for the C/A-rich GcvB target sites was determined by MEME (http://meme.nbcr.net/meme4 1/cgi-bin/meme.cgi).

C) *(Left) E. coli* Δ*gcvB recA*$^-$ strains carrying the *ilvC::gfp*, *iciA::gfp*, *ilvE::gfp or thrL::gfp* fusion plasmids in combination with control vector pTP11 (black), or plasmids expressing *Salmonella* wild-type GcvB RNA (pP_L*gcvB*, red) or two of the mutant alleles (pP_L*gcvB*$_{\Delta R1}$, blue; pP_L*gcvB*$_{\Delta R2}$, green) were grown to stationary phase and were subjected to flow cytometry analysis. All data acquired from the experiments are plotted in fluorescence histograms generated from all events measured (30,000 events). Cellular fluorescence is given in arbitrary units (GFP intensity). Regulation by GcvB wild-type or GcvBΔ_{R2} is visible as a shift of the fluorescence curves to the left towards lower GFP intensities. *(Right)* Proposed RNA duplexes formed by GcvB with four additional target mRNAs (*ilvC, iciA, ilvE, and thrL,).*

Figure 4: GcvB RNA represses translation of the *thrL* leader peptide.

A) Sequence of *thrLA* mRNA. The threonine leader peptide *thrL* (capital letters) is encoded upstream of the threonine biosynthetic operon *thrABC*. The transcriptional start site (Gardner, 1982), 5'UTR, coding sequence of *thrL* and region down to the ATG start codon of *thrA* are shown. Translational GFP fusions to 5th (position +15) and 19th codon (position +57, underlined sequence) were cloned. The C/A-rich threonine codons in the leader peptide are indicated by a grey bar. The predicted C/A-rich GcvB binding site upstream of the *thrL* start codon is marked by a black bar, the C/A-rich stretch that was deleted from the *thrL_19th ::gfp* fusion by a white box, respectively.

B) Western blot of *thrL_19th*, *thrL_5th* and thrL$_{\Delta CA}$::GFP fusion proteins prepared after growth to exponential phase (OD_{600nm} = 0.5) from *E. coli* Δ*gcvB recA*$^-$ carrying the indicated plasmids as in Figure 2B. GroEL was probed as loading control. Fold-changes of the different

thrL::GFP fusions upon normalization to GroEL (loading control) upon co-expression of GcvB RNAs are shown below the arrows.

Figure 5: **Regulation of the glycine/D-alanine/D-serine permease CycA by GcvB**.

A) Down-regulation of *cycA::gfp* by GcvB wild-type and mutant RNAs $GcvB_{\Delta R1}$ and $GcvB_{\Delta R2}$ after growth to stationary phase was confirmed on Western blots and by flow cytometry. Fold changes of GFP fusion protein levels on Western Blots (upon normalization to GroEL levels) by *gcvB*, $gcvB_{\Delta R1}$, or $gcvB_{\Delta R2}$ co-expression relative to the control plasmid were: -8.3/ -3.4/ -6.3. The different colours in the flow cytometric diagram indicate plasmids expressing *Salmonella* wild-type GcvB RNA ($pP_L gcvB$, red), two of the mutant alleles ($pP_L gcvB_{\Delta R1}$, blue; $pP_L gcvB_{\Delta R2}$, green), or a control plasmid (black).

B) Alignment of promoter regions and N-terminal coding sequences of *cycA* homologues in diverse enterobacteria. The transcriptional start site mapped by 5' RACE is indicated by a yellow box and the putative -10 box framed in blue, respectively.

C) GcvB inhibits 30S binding to *cycA* mRNA *in vitro*. An *in vitro* synthesized *cycA_10th::gfp* fusion mRNA fragment was used in the toeprint assay. Ribosome toeprinting of *cycA_10th::gfp* leader RNA (20 nM) as described in Experimental Procedures. '+/-' indicate the presence or absence of 30S subunit (200 nM) and fMet initiator tRNA (1 μM). The *cycA::gfp* AUG start codon position is shown. The arrow indicates the 30S toeprint. Increasing concentrations of GcvB RNA (lanes 4-6: 20, 100 and 200 nM) in the reactions inhibit 30S binding, whereas the unspecific control RNA, MicA (lane 10, 200 nM) does not impair binding. Mutant RNAs $GcvB_{\Delta R1}$, $GcvB_{\Delta R2}$, and $GcvB_{3'\Delta}$ (lanes 7-9) were added at a final concentration of 200 nM. Fold-repression values for the different GcvB forms and concentrations are given below the gel.

Figure 6: *In-vitro* structure mapping of GcvB-*cycA* target mRNA complexes.

A) Identification of GcvB binding sites on *cycA* mRNA by *in vitro* probing. 5' end-labeled *cycA* RNA (~5 nM) treated with RNase T1, lead(II) (left), or RNase III (right). The synthesized *cycA* RNA fragment comprises region -163/+72 relative to the AUG start codon. The 'G' of the AUG start codon corresponds to position '+3'. (Lane C) Untreated *cycA* RNA. (Lane T1) RNase T1 ladder of hydrolysed denatured *cycA* RNA. The position of cleaved G residues is given left of the gel. (Lane OH) Alkaline ladder. *(Left)* Probing of *cycA* in the absence (lane 1 and 8) or presence of GcvB wild-type (lanes 2 and 9), $GcvB_{\Delta R1}$ (lanes 3 and 10), $GcvB_{\Delta R2}$ (lanes 4 and 11), $GcvB_{\Delta R1\&R2}$ (lanes 5 and 12), $GcvB_{\Delta 5'end}$ (lanes 6 and 13) or

$GcvB_{\Delta 3'end}$ mutant RNAs (lanes 7 and 14); final concentrations in all lanes: ~500nM. Protected regions in *cycA* mRNA are indicated by red and blue bars and RNase III cleavage sites *(Right*) in the same regions by red arrows. The blue arrows denote additional specific RNase III cleavage sites of *cycA* RNA in the presence of GcvB wild-type or mutant RNAs.

B) Identification of *cycA* binding sites on GcvB RNA by *in vitro* probing. 5' end-labeled GcvB RNA (~5 nM) was subjected to RNase T1, lead(II) (left) in the absence (lanes 1 and 3) or presence of cold *cycA* RNAs (final concentration in lanes 2 and 4: ~500 nM). Duplexes were confirmed by RNaseIII cleavages (right) in the absence (lanes 1) or presence of cold *cycA* (final concentration in lane 2: ~500 nM). (Lane C) Untreated GcvB RNA. (Lane T1) RNase T1 ladder of hydrolysed denatured GcvB RNA. The position of cleaved G residues is given left of the gel. (Lane OH) Alkaline ladder. Protected regions are indicated by red and blue vertical bars and RNase III cleavage sites by red and blue arrows, respectively. The approximate positions of stem-loop structures SL1 and SL2 according to the GcvB RNA structure shown in C are indicated to the right of the gel.

C) Two proposed interaction sites of GcvB-*cycA* complexes. SD and AUG start codon sequences of *cycA* mRNA are boxed. In addition, the G/U-rich consensus R1 and consensus R2 are boxed in GcvB RNA. The coloured residues (blue invoving R1, red involving R2) were protected from lead(II) cleavage upon duplex formation (see B).

Figure 7: **The GcvB network.** GcvB sRNA represses mRNAs of the periplasmic substrate proteins of many ABC transporters involved in amino acid uptake. Seven targets which were validated by *in vitro* structure probing in our previous study (Sharma et al., 2007) are shown as grey ovals; blue ovals denote additional proteins which are controlled by GcvB and were identified in this study from transcriptomic analysis or bioinformatic-based target predictions. The newly identified genes include several genes involved in transport systems or amino acid biosynthesis. The preferred amino acid, di- or oligopeptide substrates of relevant periplasmic binding proteins are shown in yellow ovals. Control of the *gcvB* gene by two transcription factors (GcvR, GcvA) of the glycine cleavage system in *E. coli* is based on extensive work by the Stauffer laboratory (see text for references).

Figure 8: **Analysis of translational repression by GcvB RNAs in an *in-vitro* translation system.** *In vitro* synthesized, full-length mRNAs (100 nM) of *lrp::3XFLAG* and the *cycA_19th::gfp*, *gltI_23rd::gfp*, and *ompD_39th::gfp* fusions were *in vitro* translated with

reconstituted 70S ribosomes in the presence of 100 nM Hfq as described in Materials and Methods. Syntheses of Lrp::3XFLAG and GFP fusion protein levels were determined after incubation for 30 min by Western blot analysis. Where indicated, the reactions contained a 10-fold excess of GcvB WT or mutant RNAs as indicated above the lanes.

EXPERIMENTAL PROCEDURES

Bacterial strains, oligonucleotides and plasmids

The *S.* Typhimurium and *E. coli* strains used in this section are listed in Supplementary Table S4. The complete list of DNA oligonucleotides used for cloning, PCR amplification of T7 templates, toeprinting assays, and as hybridization probes and as probes in hybridization is provided in Table S5. Plasmids that were used or constructed in this study are listed in Supplementary Table 6. Details of their construction, and insert sequences are given in Supplementary Tables 7-10.Plasmid pBAD-GcvB (pKP1-1) was constructed as described for pBAD-RybB in (Papenfort et al., 2006), but using primers JVO-0897 and pZE-XbaI for insert amplification on pTP05 which carries the *Salmonella gcvB* locus (292 bp upstream of the +1 site to 116 bp downstream of the terminator). Similarly, plasmid pJL3-15 and primers JVO-0897 and pZE-XbaI were used for amplification of the $gcvB_{\Delta R1}$ allele which was cloned under the control of the pBAD promoter analogous to pKP1-1 resulting in plasmid pKP2-6 (pBAD-GcvB$_{\Delta R1}$). Plasmid pKP30-1 (pBAD-GcvB$_{\Delta R2}$) was constructed by PCR amplification from plasmid pKP1-1 using Phusion Polymerase and primer pair JVO-0895/JVO-0896, DpnI digestion of template plasmid, and self-ligation of purified PCR product.

GcvB mutant plasmids were constructed via PCR amplification from the original plasmids pTP05 or pTP09 using Phusion Polymerase (Finnzymes, #F-530S), DpnI digestion of template plasmid, and self-ligation of purified PCR products. Sequences of the mutant inserts are given in Supplementary Table 7. Consensus R2 was deleted from plasmid pP_L $gcvB_{\Delta R1}$ (pJL22) by PCR amplification of pJL22 using oligos JVO-0895/-0896 and Phusion Polymerase, DpnI digestion and self-ligation of purified PCR product yielding plasmid pP_L $gcvB_{\Delta R1\& \Delta R2}$ (pJL36-5) in which nucleotides 66 – 89 and 136 – 144 of GcvB are deleted. The promoterless GcvB plasmid p$gcvB_{\Delta PL}$ (pJL85-4) was constructed by PCR amplification of plasmid pTP09 using oligos JVO-3355/-1396 and Phusion Polymerase, DpnI digestion and self-ligation of purified PCR product. In this plasmid the P_{LacO} promoter region from (-1) to (-35) according to the GcvB transcriptional start site is deleted.

Translational GFP fusions to GcvB target mRNAs were constructed as described (Urban & Vogel, 2007). The 5'UTR parts included in the fusion plasmids start at transcriptional start sites derived from own 5'RACE results, promoters described in the EcoCyc (http://biocyc.org/) database or the literature, or deep-sequencing results of Hfq-bound RNA in *Salmonella* (Sittka et al., 2008).

For *cycA*, the promoter was mapped by 5' RACE. Afterwards, the BseRI/NheI-digested cDNA product was cloned into the BsgI/NheI-digested fusion plasmid pXG-20 (Urban & Vogel, 2007). All other fusions were cloned into vector pXG-10. The details for GFP plasmid construction are listed in Supplementary Table 8, and the inserts of all GFP fusions are listed in Supplementary Table 9.

A shortened *cycA::gfp* fusion to the 10th amino acid was constructed by PCR amplification from the original plasmid pJL30-14 using Phusion polymerase and oligos JVO-3330 and JVO-0323, DpnI digestion of template plasmid, NheI digestion, and self-ligation of purified PCR products. This resulted in plasmid pJL83-2.

For the verification of regulation of *thrL* by GcvB, the following GFP fusions were constructed: *thrL_19th::gfp* (fusion to 19th amino acid) and a shorter version *thrL_5th::gfp* (fusion to 5th amino acid) as well as a $thrL_{\Delta CA}$*::gfp* (deletion of seven nucleotides ((-4) to (-10) according to the start codon) from the predicted C/A-rich GcvB binding site from the *thrL_19th::gfp* fusion). Further details of the different *thrL*::gfp fusions are listed in Supplementary Table 6. Shorter version of *thrL* (*thrL_5h*) and $thrL_{\Delta CA}$ were constructed by PCR amplification from the original plasmid of the *thrL_19th* fusion pSP20-1 using Phusion polymerase and the oligos JVO-0323/JVO-5951 or CSO-0001/CSO-0002, respectively. PCRs were followed by DpnI digest of the template. For the shorter version of *thrL::gfp* (*thrL_5th*), the purified PCR product was NheI-digested, followed by ligation, which resulted in pSP37-1. Self-ligation of PCR-products for $thrL_{\Delta CA}$ version leads to pSP-38-3.

The *Salmonella SL1344 lrp gene* was C-terminally FLAG-tagged by using primers JVO-4981 and JVO-4982 and a modified lambda red approach (Uzzau *et al.*, 2001), resulting in *lrp*::3xFLAG (strain JVS-5412). Afterwards, the *lrp*::3xFLAG fusion was transduced in a *Salmonella* wild-type strain (JVS-1574), resulting in JVS-5417.

Media and growth conditions

Bacterial cells were grown aerobically at 37°C in Lennox broth or M9 minimal medium supplemented with 0.4 % glucose. When required, antibiotics were added at 92 µg/µl streptomycin, 100 µg/ml ampicillin, 50 µg/ml kanamycin, and 20 µg/ml chloramphenicol (final concentrations).

Transcriptomic experiments

Salmonella RNA was isolated using the Promega SV total RNA purification kit. Microarrays used in this study were produced by *in-situ* synthesis as 8x15k multipack format from Agilent Technologies (Agilent AMADID Designcode:026881). Each microarray comprises 13268 60-mer *S. Typhimurium* strain SL1344-specific oligonucleotides supplemented with 360 60-mer oligonucleotides specific for 149 Salmonella sRNAs. The experimental design involves the use of *S.* Typhimurium genomic DNA as the co-hybridized control for one channel on all microarrays. Total RNA and chromosomal DNA were labelled by random priming according to the protocols described at the IFR (Institute of Food Research, Norwich) website (www.ifr.ac.uk/safety/microarrays/protocols.html). Briefly, 10 µg RNA were reverse transcribed and labeled with Cy5-conjugated dCTP (Pharmacia) using 200 units of StrataScript and random hexamers (Invitrogen). Chromosomal DNA (400 ng) was labeled with Cy3-dCTP using the Klenow fragment. After purificaion, each Cy5-labelled cDNA sample was combined with Cy3-labelled chromosomal DNA and hybridized to a microarray overnight at 65°C using hybridization and blocking buffer according the suppliers' instructions. After hybridization, slides were washed with a standard SSC protocol and scanned at 5 µm and extended range using a G2565CA high resolution laser microarray scanner (Agilent Technologies). Raw microarray image data were analyzed with the Image Analysis/Feature Extraction software G2567AA (Version A.10.5.1, Agilent Technologies). For each array feature both MedianSignals were background corrected by appropriate BGMedianSignal subtraction while signals <10 were adjusted to 10. To compensate for unequal dye incorporation, data centering to zero was performed for each single microarray on one slide. Microarray data were analysed using GeneSpring 7.3 (Agilent) and genes were considered to be differentially expressed if they displayed ≥2-fold changes in both replicates and were statistically significantly different (Student's t-test; $p<0.1$). The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE26573 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26573).

RNA and protein detection

RNA preparation and Northern analysis as well as GFP fusion and GroEL protein detection followed previously published protocols (Urban & Vogel, 2007). GcvB RNAs were detected with 5' end-labeled oligos (JV0-0749), and 5S rRNA with oligo JVO-0322, respectively. In

addition, *lrp* mRNA and *ndk* mRNAs were detected with 5' end-labeled oligos JV0-2978 and JVO-1493, respectively.

For flow cytometric analysis, *E. coli* strains carrying *gfp* fusion plasmids were grown to stationary phase in 3 ml liquid cultures (inoculated from single colonies) in LB medium (supplemented with appropriate antibiotics). After 14 h, bacteria from 1 ml culture volume were harvested by centrifugation for 2 min at 8,000 rpm and 4°C. After removal of supernatants, bacteria were resuspended in 500 µl 2% PFA (paraformaldehyde) in 1×PBS buffer (pH 7.4) for fixation and stored for up to five days in the dark at 4°C until flow cytometric analysis. Prior to FACS measurements, samples were diluted 1:250 in 1× PBS (phosphate-buffered saline) buffer, pH 7.4 buffer.

To determine reporter activities of single cells, a BD FACSCantoTM Flow Cytometer equipped with a blue excitation source (air-cooled, 20 mW solid state 488 nm laser) was used to measure forward angle light scatter (FSC), side scatter (SSC), and the fluorescence of the cells (FITC). The instrument settings were in logarithmic mode: FSC-H: 516, SSC-A: 626; FITC-A (GFP): 753. GFP fluorescence intensity was measured for 30,000 events (maximum threshold of 10,000 events/sec). All flow cytometric analyses were done in duplicates. Data analysis was carried out using FCS Express software, version 3 (De Novo Software).

Colony fluorescence imaging

E. coli carrying *gfp* fusion plasmids were grown on LB plates over-night. Colonies were photographed with a FUJI LAS-3000 image analyzer with a 510 nm emission filter and excitation at 460 nm.

In vitro *structure probing and 30S ribosome toeprinting*

DNA templates carrying a T7 promoter sequence for *in vitro* transcription were generated by PCR. Primers and sequences of the T7 transcripts are included in Supplementary Tables 10 and 11. RNA was *in vitro* transcribed, quality-checked and labeled at the 5' end as described (Sittka *et al.*, 2007). The protocol for 5' end labeling of RNA is published in (Papenfort et al., 2006). Secondary structure probing and mapping of RNA complexes was conducted on 5'-end-labelled RNA (~0.1 pmol) in 10 µl reactions as previously described (Sharma et al., 2007). In brief, RNA was denatured 1 min at 95°C and chilled on ice for 5

min, upon which 1 µg yeast RNA and 10X structure buffer (0.1 M Tris pH 7, 1 M KCl, 0.1 M $MgCl_2$, Ambion) were added. Concentrations of unlabelled sRNA/mRNA leader added to the reactions are given in the figure legends. Following incubation for 10 min at 37°C, 2 µl of a fresh solution of lead(II) acetate (25 mM; Fluka #15319), or 2 µl of RNase T1 (0.05 u/µl; Ambion, #AM2283) were added and incubated for 2 or 3 min at 37°C, respectively. RNase III cleavage reactions contained 1 mM DTT and 1.3 unit enzyme (NEB #M0245S), and were incubated for 6 min at 37 °C.

Reactions were stopped by addition of 12 µl loading buffer (95% Formamide, 18mM EDTA, and 0.025% SDS, Xylene Cyanol, and Bromophenol Blue, Ambion) on ice. RNase T1 ladders were obtained by incubating labeled RNA (~0.2 pmol) in 1X sequencing buffer (Ambion) for 1 min at 95°C. Subsequent, 1µl RNase T1 (0.1 u/µl) was added and incubation continued at 37°C for 5 min. OH ladders were generated by 5 min incubation of 0.2 pmol labelled RNA in alkaline hydrolysis buffer (Ambion) at 95°C. Reactions were stopped with 12 µl loading buffer. Samples were denatured for 3 min at 95°C prior to separation on 6% polyacrylamide/7M urea sequencing gels in 1X TBE. Gels were dried analyzed using a PhosphorImager (FLA-3000 Series, Fuji), and AIDA software (Raytest, Germany).

Toeprinting reactions were carried out as described (Sharma et al., 2007). Specifically, 0.2 pmol of an unlabelled *cycA 10th::gfp* mRNA fragment (171 nt, T7 template amplified with JVO-1274/-1976) and 0.5 pmol of 5'end labelled primer JVO-1976 complementary to the *gfp* coding region were annealed. For inhibition analysis, 0.2, 1 and 2 pmol of GcvB RNA or 2 pmol control RNA (MicA) or GcvB mutant RNAs were added.

In vitro translation assays

Translation reactions were carried out with PureSystem (Cosmo Bio Co., Ltd, PGM-PURE2048C) according to the manufacturer's instructions. In brief, 100 pmol *in vitro* transcribed mRNA (*lrp::*3xFLAG, *cycA::gfp*, *gltI::gfp*, *ompD::gfp*; see Supplementary Table S10 and S11) were denatured in absence or presence of 1000 pmol GcvB, GcvBΔR1, GcvBΔR2, or GcvBΔR1&R2 RNA for 1 min at 95 °C and chilled for 5 minutes on ice. Hfq (100 pmol) was mixed with mRNA (and sRNA) and pre-incubated for 10 min at 37°C before addition of PureSystem mix. Translation was performed in a 10 µl reaction for 30 min at 37°C, stopped with four volumes of ice-cold acetone and chilled on ice for 15 min. Proteins were collected by centrifugation (10,000*g*, 10 min, 4°C) and quantified by Western blot analysis with monoclonal FLAG or GFP antibody (Urban & Vogel, 2007). Ribosomes were

detected via S1 antibody (1:5000, kindly provided by M. Springer, IBPC Paris, France) and anti-rabbit secondary antibody (conjugated with horseradish peroxidase; GE Healthcare).

Motif detection using MEME and MAST

The *gfp* fusion sequences of the seven previously-defined targets (*dppA, oppA, livJ, livK, argT*, STM4351, and *gltI*, see Supplementary Table 9) and nine new targets from the transcriptomic analysis (*yaeC*, *serA*, *ybdH*, *brnQ*, *tppB, ygjU(sstT) ndk, gdhA,* and *lrp*, see Supplementary Table 9) were used as input sequences for MEME (http://meme.sdsc.edu/meme/cgi-bin/meme.cgi) motif identification (Bailey & Gribskov, 1998). The following parameters were defined: number of different motifs: 3; minimum number of sites: 10; minimum motif width: 6; maximum motif width: 25. The position-specific weight matrix for the C/A-rich GcvB target derived from the MEME search was then used as input for MAST (http://meme.sdsc.edu/meme/cgi-bin/mast.cgi) (Bailey & Gribskov, 1998) searches in a database composed of the 5' regions (-70 to +30 according to the start codon) of all annotated *Salmonella* ORFs (Supplementary Table 2). These were extracted as a multi-Fasta file from the *Salmonella Typhimurium* LT2 genome sequence and annotation (NC_003197) using own unpublished Perl *(*www.perl.org/) scripts.

Prediction of sRNA-target mRNA duplexes

GcvB-target mRNA complexes were predicted with TargetRNA (Tjaden, 2008). An extended GcvB consensus region R1 (TTGGCTTACGGTTGTGATGTTGTGTTGTTGTGTTTGCAATTGGTCTGCG) was used as small RNA input and interactions with the 5' regions (-70/+30 nt of annotated start codon) of all *Salmonella* LT2 genes with default settings (at least 9 bp hybridization seed and no G:U pairs in seed) were predicted. Only interactions with a p-value cut-off <= 0.05 were further considered (see Supplementary Table 3). Subsequently, we investigated which of the predicted targets with a p-value below 0.05 show more than 2-fold change in the transcriptomic experiments and which of the targets contain the C/A-rich motif (Fig. 3B), that was derived from the 16 targets from our previous (Sharma et al., 2007) and the transcriptomic experiments in this study (Supplementary Table 2 & 3).

Table 1: The GcvB regulon of *S.* Typhimurium contains at least 45 mRNAs. The values correspond to the fold-changes of particular genes in the indicated strains compared to the expression level in the *Salmonella ΔgcvB* strain harbouring the pBAD-control plasmid ('Δ*gcvB* + Ctr'). A negative value corresponds to down-regulation and a positive value to up-regulation of the mRNA, respectively. Columns number 5-7 indicate the fold-changes for the genes given in the first column in three microarray experiments, each of which was carried out in biological duplicate, and involved wild-type or mutant GcvB RNAs. A p-value < 0.1 was used as cut-off. Genes that show an R1-dependent pattern of down-regulation, i.e. more than 2-fold change (p-value < 0.1) the two strains 'Δ*gcvB* + GcvB' and 'Δ*gcvB* + ΔR2' are marked in grey. Operons are indicated in bold.

Gene	Description	Amino acid transport	Amino acid metabolism/biosynthesis	Fold-change (*vs. ΔgcvB* + pBAD-Ctr)			Hfq binding		Regulation			Regulated in E. coli[4]
				ΔgcvB + GcvB WT	*ΔgcvB* + GcvB R1	*ΔgcvB* + GcvB R2	Enriched in Hfq-coIP OD 2.0[2]	Enriched in Hfq-coIP over growth[3]	Plate	WB	FACS	
gcvB	GcvB sRNA			>100	>100	>100	yes					
argT	ABC superfamily (binding_protein)	+		-2.7	-1.2	-2.3		1	no[1,#]	yes[1]	n.d.	MA
aroP	aromatic amino acid transport protein	+		-3.0	-1.3	-3.2						
astA	arginine succinyltransferase		+	-3.6	±1.0	-4.0						
astC (argD)	succinylornithine transaminase		+	-3.8	-1.0	-3.6						
brnQ	branched-chain amino acid transport system II carrier protein (liv-II)	+		-4.8	-1.3	-5.2			no[#]	yes	no	
citA	citrate-proton symporter			-2.1	+1.1	-2.9						
cybC	cytochrome b(562)			-2.1	-2.1	+1.4						
cycA	APC family D-alanine/D-serine/glycine transport protein	+		-5.1	-4.6	-4.9		3	yes	yes	yes	yes
dlhH	putative dienelactone hydrolase			-2.5	±1.0	-2.6	yes					
dppA	ABC superfamily (peri_perm); dipeptide transport protein	+		-29.3	-2.3	-33.2		1	yes[1]	yes[1]	n.d.	yes

dppB	ABC superfamily (membrane); dipeptide transport system permease protein 1	+		-17.7	-1.8	-19.0						MA
dppD	ABC superfamily (atp_bind); putative ATP-binding component of dipeptide transport system	+		-15.7	-1.4	-16.9						MA
ecnB	entericidin B membrane lipoprotein			-4.5	-1.2	-4.7	yes	3				
gdhA	NADP-specific glutamate dehydrogenase		+	-6.1	-3.2	-5.7			yes	yes	yes	MA
gltI	ABC superfamily (bind_prot); putative periplasmic binding transport protein. glutamate/aspartate transporter	+		-5.5	+1.2	-5.9	yes	1	yes	n.d.	n.d.	
gltJ	ABC superfamily (membrane); glutamate/aspartate transport system permease	+		-10.3	-2.0	-8.5						
gltK	glutamate/aspartate transporter	+		-9.6	-2.1	-8.5		2				
gltL	ABC superfamily (atp_bind); ATP-binding protein of glutamate/aspartate transport system	+		-4.7	-1.8	-4.8						
ilvN	acetolactate synthase I. valine sensitive. small subunit			-2.2	+1.1	-2.2						MA
livJ	ABC superfamily (bind_prot); high-affinity branched-chain amino acid transporter	+		-3.4	-1.8	-3.8			yes[1]	yes[1]	n.d.	
lrp	leucine-responsive regulatory protein		+	-2.0	-2.5	-1.8	yes		yes	no	yes	
ndk	nucleoside diphosphate kinase			-3.5	+1.1	-3.6	yes	1	yes	no	yes	
oppA	ABC superfamily (periplasm); periplasmic oligopeptide-binding protein precursor	+		-14.6	+1.1	-14.0	yes	2	yes	yes	yes	
oppB	ABC superfamily (membrane); oligopeptide transport system permease protein	+		-12.7	+1.1	-11.8	yes	1				
oppC	ABC superfamily (membrane); oligopeptide transport system permease protein	+		-5.5	-1.0	-5.0	yes					
oppD	ABC superfamily (atp-binding); oligopeptide transport ATP-binding protein	+		-3.9	+1.1	-3.2	yes					
oppF	ABC superfamily (atp-binding); oligopeptide transport ATP-binding protein	+		-2.8	+1.1	-3.1	yes					
pepN	aminopeptidase N		+	-2.4	±1.0	-2.6	yes					
pheA	chorismate mutase-P and prephenate dehydratase		+	-2.6	+1.1	-3.0						
phnA	putative alkylphosphonate uptake protein in phosphonate metabolism		+	+2.0	+1.4	+2.0		2				
purU	formyltetrahydrofolate deformylase			-3.8	+1.1	-3.6						
putP	SSS family major sodium/proline symporter (proline permease)	+		-3.1	-1.2	-3.2						
serA	D-3-phosphoglycerate dehydrogenase		+	-2.4	-1.4	-2.7			n.d.[#]	yes	yes	MA
sfbA	putative lipoprotein. putative ABC-type transport system	+		-3.2	+1.2	-3.9						

	ATPase component/cell division protein											
sfbB	putative ABC-type transport system ATPase component/cell division protein	+		-3.0	+1.2	-3.2						
spy	periplasmic protein related to spheroblast formation			-2.4	-1.2	-2.7						
trpE	anthranilate synthase component I		+	-2.1	±1.0	-2.1						
yaeC (metQ)	putative outer membrane lipoprotein. DL-methionine transporter substrate-binding subunit	+		-2.5	±1.0	-2.1			yes	no	yes	
ybdH	putative glycerol dehydrogenase			-3.9	-1.3	-3.6			no[#]	n.d.	yes	
ydeJ	competence damage-inducible protein A			-2.0	-1.3	-2.2						
tppB	putative POT family peptide transport protein. putative tripeptide transporter permease	+		-7.8	-1.0	-7.9		1	yes	no	yes	
ydiJ	putative oxidase			-2.7	-1.1	-3.0						
yecS	putative ABC-type amino acid transporter	+		-2.8	-1.0	-2.9						
yedO	putative 1-cyclopropane-carboxylate deaminase. D-cysteine desulfhydrase		+	-3.0	-1.0	-2.8						
ygjU (sstT)	serine/threonine transporter SstT	+		-10.2	+1.2	-10.6	yes	3	yes	yes	yes	yes
yicL	putative permease			-2.1	-1.0	-2.2						
yifK	putative APC family amino-acid transport protein	+		-6.3	-1.0	-6.1			##	##	##	MA
STM1494	putative transport system permease protein; ABC-type transport systems			-3.1	-1.2	-3.2						n.c.
STM1638	putative SAM-dependent methyltransferases			-5.2	+1.1	-5.0	yes					n.c.
STM1747	putative inner membrane protein			-4.9	±1.0	-5.4						n.c.
STM4195	putative Na+-dependent transporter			-3.1	+1.2	-3.7						n.c.
STM4351	putative arginine-binding periplasmic protein	+		-3.0	+1.2	-3.3			no[1,#]	yes[1]	n.d.	n.c.
STnc310	ncRNA			-3.0	-1.0	-2.8						n.c.
traT	conjugative transfer: surface exclusion			-2.2	-2.0	-1.9						n.c.
Additional targets based on biocomputational predictions												
ilvC	ketol-acid reductoisomerase		+	-1.2	1.0	-1.5			yes	yes	yes	MA
thrL	*thr* operon leader peptide		+	1.8	-1.3	1.4		2	yes	yes	yes	
iciA (argP)	transcriptional activator, chromosome replication initiation inhibitor		+	-1.6	1.1	-1.7			yes	yes	yes	
ilvE	branched-chain amino acid aminotransferase		+	-1.8	1.1	-1.9			yes	n.d.	yes	
livK	high-affinity branched-chain amino acid transporter	+		-1.5	-1.5	-1.7						MA

[1] Previously determined in (Sharma et al., 2007).
[2] Hfq-binding at OD600 = 2 was determined by coIP of *Salmonella* Hfq::3XFLAG combined with RNA-seq (Sittka et al., 2008).
[3] Chao & Vogel, unpublished results. A coIP of *Salmonella* Hfq::3xFLAG combined with RNA-seq as previously described (Sittka et al., 2008) was performed from seven different samples during growth. Specifically, bacteria were grown in LB broth and samples were harvested at early exponential, mid-exponential, late-exponential, OD600 = 2, OD600 = 2+3 hrs, OD600 = 2+6 hrs, and growth overnight. Numbers indicate the number of growth conditions under which the mRNA was detected and an enrichment in the Hfq-coIP was observed compared to the control condition.
[4] Regulation was either observed on microarrays (MA) in *E. coli* or was validated by reporter gene fusions (Urbanowski et al., 2000, Pulvermacher *et al.*, 2009a, Pulvermacher et al., 2009b, Pulvermacher et al., 2009c). "n.c." indicates *Salmonella* specific genes.

[#] Low fluorescence
[##] GFP fusion was not fluorescent
n.d.: not determined

REFERENCES

Ansong, C., H. Yoon, S. Porwollik, H. Mottaz-Brewer, B. O. Petritis, N. Jaitly, J. N. Adkins, M. McClelland, F. Heffron & R. D. Smith, (2009) Global systems-level analysis of Hfq and SmpB deletion mutants in Salmonella: implications for virulence and global protein translation. *PLoS One* **4**: e4809.

Antal, M., V. Bordeau, V. Douchin & B. Felden, (2005) A small bacterial RNA regulates a putative ABC transporter. *J Biol Chem* **280**: 7901-7908.

Argaman, L. & S. Altuvia, (2000) fhlA repression by OxyS RNA: kissing complex formation at two sites results in a stable antisense-target RNA complex. *J Mol Biol* **300**: 1101-1112.

Argaman, L., R. Hershberg, J. Vogel, G. Bejerano, E. G. Wagner, H. Margalit & S. Altuvia, (2001) Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. *Curr Biol* **11**: 941-950.

Bailey, T. L. & M. Gribskov, (1998) Methods and statistics for combining motif match scores. *J Comput Biol* **5**: 211-221.

Balbontin, R., F. Fiorini, N. Figueroa-Bossi, J. Casadesus & L. Bossi, (2010) Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in Salmonella enterica. *Mol Microbiol* **78**: 380-394.

Beisel, C. & G. Storz, (2011) The base pairing RNA Spot42 participates in a multi-output feedforward loop to help enact catabolite repression in Escherichia coli. *Molecular Cell* **41**: 286-297.

Bohn, C., C. Rigoulay, S. Chabelskaya, C. M. Sharma, A. Marchais, P. Skorski, E. Borezee-Durant, R. Barbet, E. Jacquet, A. Jacq, D. Gautheret, B. Felden, J. Vogel & P. Bouloc, (2010) Experimental discovery of small RNAs in Staphylococcus aureus reveals a riboregulator of central metabolism. *Nucleic Acids Res* **38**: 6620-6636.

Boisset, S., T. Geissmann, E. Huntzinger, P. Fechter, N. Bendridi, M. Possedko, C. Chevalier, A. C. Helfer, Y. Benito, A. Jacquier, C. Gaspin, F. Vandenesch & P. Romby, (2007) Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* **21**: 1353-1366.

Bouvier, M., C. M. Sharma, F. Mika, K. H. Nierhaus & J. Vogel, (2008) Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* **32**: 827-837.

Boysen, A., J. Moller-Jensen, B. Kallipolitis, P. Valentin-Hansen & M. Overgaard, Translational regulation of gene expression by an anaerobically induced small non-coding RNA in Escherichia coli. *J Biol Chem* **285**: 10690-10702.

Busch, A., A. S. Richter & R. Backofen, (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**: 2849-2856.

Busi, F., J. Le Derout, M. Cerciat, P. Regnier & E. Hajnsdorf, (2010) Is the secondary putative RNA-RNA interaction site relevant to GcvB mediated regulation of oppA mRNA in Escherichia coli? *Biochimie*.

Calvo, J. M. & R. G. Matthews, (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. *Microbiol Rev* **58**: 466-490.

Chabelskaya, S., O. Gaillot & B. Felden, (2010) A Staphylococcus aureus small RNA is required for bacterial virulence and regulates the expression of an immune-evasion molecule. *PLoS Pathog* **6**: e1000927.

Chen, C. F., J. Lan, M. Korovine, Z. Q. Shao, L. Tao, J. Zhang & E. B. Newman, (1997) Metabolic regulation of lrp gene expression in Escherichia coli K-12. *Microbiology* **143 (Pt 6)**: 2079-2084.

Chevalier, C., T. Geissmann, A. C. Helfer & P. Romby, (2009) Probing mRNA structure and sRNA-mRNA interactions in bacteria using enzymes and lead(II). *Methods Mol Biol* **540**: 215-232.

Cosloy, S. D., (1973) D-serine transport system in Escherichia coli K-12. *J Bacteriol* **114**: 679-684.

Desnoyers, G., A. Morissette, K. Prevost & E. Masse, (2009) Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA. *EMBO J* **28**: 1551-1561.

Durand, S. & G. Storz, (2010) Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol Microbiol* **75**: 1215-1231.

Gardner, J. F., (1982) Initiation, pausing, and termination of transcription in the threonine operon regulatory region of Escherichia coli. *J Biol Chem* **257**: 3896-3904.

Ghrist, A. C. & G. V. Stauffer, (1995) The Escherichia coli glycine transport system and its role in the regulation of the glycine cleavage enzyme system. *Microbiology* **141 (Pt 1)**: 133-140.
Gottesman, S. & G. Storz, (2011) Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol*.
Guillier, M. & S. Gottesman, (2008) The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res* **36**: 6781-6794.
Guillier, M., S. Gottesman & G. Storz, (2006) Modulating the outer membrane with small RNAs. *Genes Dev* **20**: 2338-2348.
Guisbert, E., V. A. Rhodius, N. Ahuja, E. Witkin & C. A. Gross, (2007) Hfq modulates the sigmaE-mediated envelope stress response and the sigma32-mediated cytoplasmic stress response in Escherichia coli. *J Bacteriol* **189**: 1963-1973.
Hartz, D., D. S. McPheeters, R. Traut & L. Gold, (1988) Extension inhibition analysis of translation initiation complexes. *Methods Enzymol* **164**: 419-425.
Holmqvist, E., J. Reimegard, M. Sterk, N. Grantcharova, U. Romling & E. G. Wagner, (2010) Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J* **29**: 1840-1850.
Huntzinger, E., S. Boisset, C. Saveanu, Y. Benito, T. Geissmann, A. Namane, G. Lina, J. Etienne, B. Ehresmann, C. Ehresmann, A. Jacquier, F. Vandenesch & P. Romby, (2005) Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J* **24**: 824-835.
Hussein, R. & H. N. Lim, (2011) Disruption of small RNA signaling caused by competition for Hfq. *Proc Natl Acad Sci U S A* **108**: 1110-1115.
Jin, Y., R. M. Watt, A. Danchin & J. D. Huang, (2009) Small noncoding RNA GcvB is a novel regulator of acid resistance in Escherichia coli. *BMC Genomics* **10**: 165.
Johansen, J., A. A. Rasmussen, M. Overgaard & P. Valentin-Hansen, (2006) Conserved small non-coding RNAs that belong to the sigmaE regulon: role in down-regulation of outer membrane proteins. *J Mol Biol* **364**: 1-8.
Kawamoto, H., Y. Koide, T. Morita & H. Aiba, (2006) Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol Microbiol* **61**: 1013-1022.
Kikuchi, G., (1973) The glycine cleavage system: composition, reaction mechanism, and physiological significance. *Mol Cell Biochem* **1**: 169-187.
Landgraf, J. R., J. Wu & J. M. Calvo, (1996) Effects of nutrition and growth rate on Lrp levels in Escherichia coli. *J Bacteriol* **178**: 6930-6936.
Lease, R. A., M. E. Cusick & M. Belfort, (1998) Riboregulation in Escherichia coli: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc Natl Acad Sci U S A* **95**: 12456-12461.
Lynn, S. P., J. F. Gardner & W. S. Reznikoff, (1982) Attenuation regulation in the thr operon of Escherichia coli K-12: molecular cloning and transcription of the controlling region. *J Bacteriol* **152**: 363-371.
Majdalani, N., C. Cunning, D. Sledjeski, T. Elliott & S. Gottesman, (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A* **95**: 12462-12467.
Mandin, P. & S. Gottesman, Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J* **29**: 3094-3107.
Masse, E. & S. Gottesman, (2002) A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proc Natl Acad Sci U S A* **99**: 4620-4625.
Masse, E., C. K. Vanderpool & S. Gottesman, (2005) Effect of RyhB small RNA on global iron use in Escherichia coli. *J Bacteriol* **187**: 6962-6971.
McArthur, S. D., S. C. Pulvermacher & G. V. Stauffer, (2006) The Yersinia pestis gcvB gene encodes two small regulatory RNA molecules. *BMC Microbiol* **6**: 52.
McClelland, M., K. E. Sanderson, J. Spieth, S. W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston & R. K. Wilson, (2001) Complete genome sequence of Salmonella enterica serovar Typhimurium LT2. *Nature* **413**: 852-856.

Mizuno, T., M. Y. Chou & M. Inouye, (1984) A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci U S A* **81**: 1966-1970.

Novick, R. P., H. F. Ross, S. J. Projan, J. Kornblum, B. Kreiswirth & S. Moghazeh, (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* **12**: 3967-3975.

Papenfort, K., M. Bouvier, F. Mika, C. M. Sharma & J. Vogel, (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci U S A* **107**: 20435-20440.

Papenfort, K., V. Pfeiffer, F. Mika, S. Lucchini, J. C. Hinton & J. Vogel, (2006) SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. *Mol Microbiol* **62**: 1674-1688.

Papenfort, K., N. Said, T. Welsink, S. Lucchini, J. C. Hinton & J. Vogel, (2009) Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol Microbiol* **74**: 139-158.

Papenfort, K. & J. Vogel, (2009) Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res Microbiol* **160**: 278-287.

Pfeiffer, V., K. Papenfort, S. Lucchini, J. C. Hinton & J. Vogel, (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* **16**: 840-846.

Porwollik, S. & M. McClelland, (2003) Lateral gene transfer in Salmonella. *Microbes Infect* **5**: 977-989.

Prevost, K., H. Salvail, G. Desnoyers, J. F. Jacques, E. Phaneuf & E. Masse, (2007) The small RNA RyhB activates the translation of shiA mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Mol Microbiol* **64**: 1260-1273.

Pulvermacher, S. C., L. T. Stauffer & G. V. Stauffer, (2008) The role of the small regulatory RNA GcvB in GcvB/mRNA posttranscriptional regulation of oppA and dppA in Escherichia coli. *FEMS Microbiol Lett* **281**: 42-50.

Pulvermacher, S. C., L. T. Stauffer & G. V. Stauffer, (2009a) Role of the Escherichia coli Hfq protein in GcvB regulation of oppA and dppA mRNAs. *Microbiology* **155**: 115-123.

Pulvermacher, S. C., L. T. Stauffer & G. V. Stauffer, (2009b) Role of the sRNA GcvB in regulation of cycA in Escherichia coli. *Microbiology* **155**: 106-114.

Pulvermacher, S. C., L. T. Stauffer & G. V. Stauffer, (2009c) The small RNA GcvB regulates sstT mRNA expression in Escherichia coli. *J Bacteriol* **191**: 238-248.

Rice, J. B. & C. K. Vanderpool, (2011) The small RNA SgrS controls sugar–phosphate accumulation by regulating multiple PTS genes. *Nucleic Acids Res* **In press**.

Robbins, J. C. & D. L. Oxender, (1973) Transport systems for alanine, serine, and glycine in Escherichia coli K-12. *J Bacteriol* **116**: 12-18.

Russell, R. R., (1972) Mapping of a D-cycloserine resistance locus in escherichia coli K-12. *J Bacteriol* **111**: 622-624.

Salvail, H., P. Lanthier-Bourbonnais, J. M. Sobota, M. Caza, J. A. Benjamin, M. E. Mendieta, F. Lepine, C. M. Dozois, J. Imlay & E. Masse, (2010) A small RNA promotes siderophore production through transcriptional and metabolic remodeling. *Proc Natl Acad Sci U S A* **107**: 15223-15228.

Sharma, C. M., F. Darfeuille, T. H. Plantinga & J. Vogel, (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev* **21**: 2804-2817.

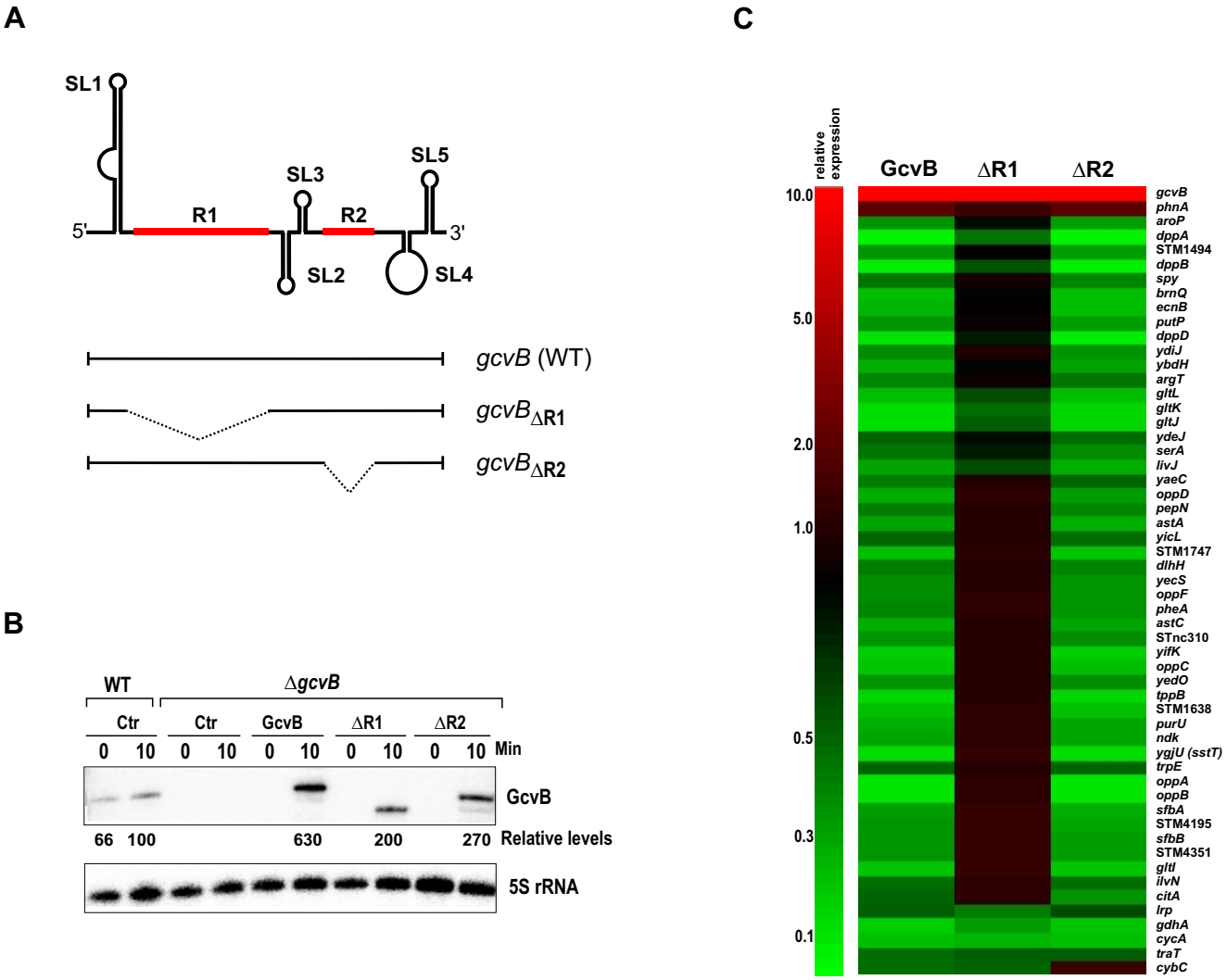
Silveira, A. C., K. L. Robertson, B. Lin, Z. Wang, G. J. Vora, A. T. Vasconcelos & F. L. Thompson, Identification of non-coding RNAs in environmental vibrios. *Microbiology* **156**: 2452-2458.

Sittka, A., S. Lucchini, K. Papenfort, C. M. Sharma, K. Rolle, T. T. Binnewies, J. C. Hinton & J. Vogel, (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* **4**: e1000163.

Sittka, A., V. Pfeiffer, K. Tedin & J. Vogel, (2007) The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Molecular microbiology* **63**: 193-217.

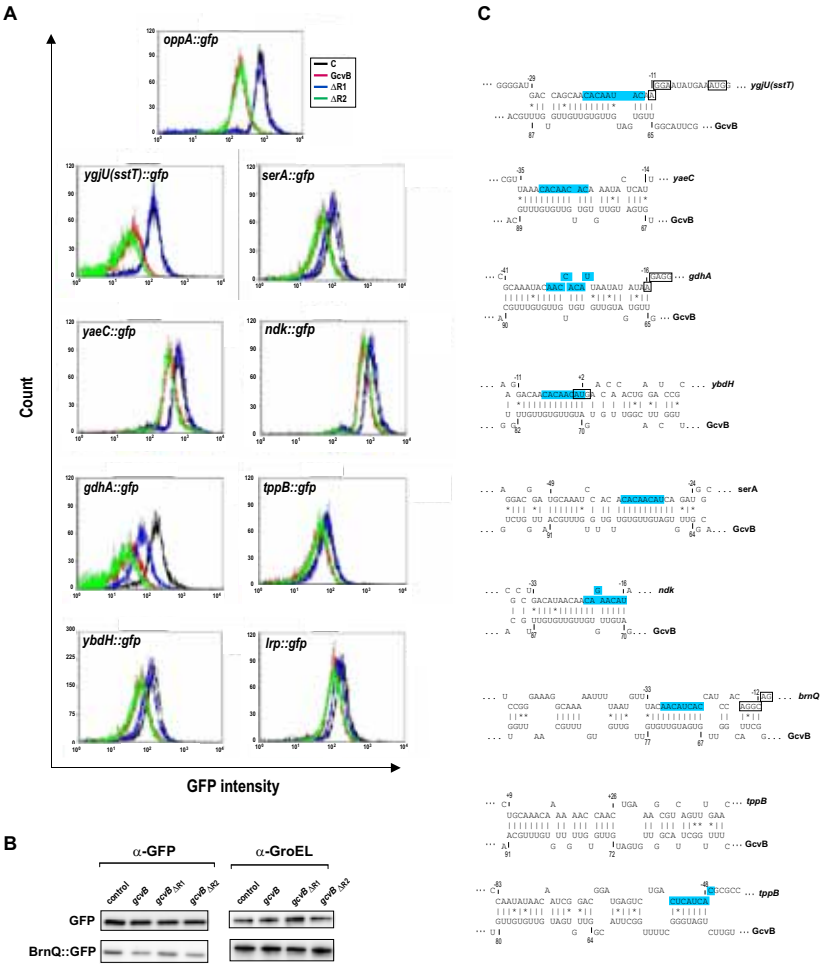
Sittka, A., C. M. Sharma, K. Rolle & J. Vogel, (2009) Deep sequencing of Salmonella RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes. *RNA Biol* **6**: 266-275.
Stauffer, L. T. & G. V. Stauffer, (1994) Characterization of the gcv control region from Escherichia coli. *J Bacteriol* **176**: 6159-6164.
Stauffer, L. T. & G. V. Stauffer, (2005) GcvA interacts with both the alpha and sigma subunits of RNA polymerase to activate the Escherichia coli gcvB gene and the gcvTHP operon. *FEMS Microbiol Lett* **242**: 333-338.
Tani, T. H., A. Khodursky, R. M. Blumenthal, P. O. Brown & R. G. Matthews, (2002) Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci U S A* **99**: 13471-13476.
Tjaden, B., (2008) TargetRNA: a tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res* **36**: W109-113.
Tjaden, B., S. S. Goodwin, J. A. Opdyke, M. Guillier, D. X. Fu, S. Gottesman & G. Storz, (2006) Target prediction for small, noncoding RNAs in bacteria. *Nucleic Acids Res* **34**: 2791-2802.
Urban, J. H. & J. Vogel, (2007) Translational control and target recognition by Escherichia coli small RNAs in vivo. *Nucleic Acids Res* **35**: 1018-1037.
Urbanowski, M. L., L. T. Stauffer & G. V. Stauffer, (2000) The gcvB gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in Escherichia coli. *Mol Microbiol* **37**: 856-868.
Uzzau, S., N. Figueroa-Bossi, S. Rubino & L. Bossi, (2001) Epitope tagging of chromosomal genes in Salmonella. *Proc Natl Acad Sci U S A* **98**: 15264-15269.
Valentin-Hansen, P., M. Eriksen & C. Udesen, (2004) The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol* **51**: 1525-1533.
Vanderpool, C. K. & S. Gottesman, (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* **54**: 1076-1089.
Vecerek, B., I. Moll & U. Blasi, (2007) Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *EMBO J* **26**: 965-975.
Vogel, J. & K. Papenfort, (2006) Small non-coding RNAs and the bacterial outer membrane. *Curr Opin Microbiol* **9**: 605-611.
Wargel, R. J., C. A. Hadur & F. C. Neuhaus, (1971) Mechanism of D-cycloserine action: transport mutants for D-alanine, D-cycloserine, and glycine. *J Bacteriol* **105**: 1028-1035.
Wilms, I., B. Voss, W. R. Hess, L. Leichert & F. Narberhaus, (2011) Small RNA-medaited control of Agrobacterium tumefaciens GABA binding protein. *Molecular Microbiology* **In press**.
Zhang, A., S. Altuvia, A. Tiwari, L. Argaman, R. Hengge-Aronis & G. Storz, (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *EMBO J* **17**: 6061-6068.
Zhang, A., K. M. Wassarman, C. Rosenow, B. C. Tjaden, G. Storz & S. Gottesman, (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* **50**: 1111-1124.

Sharma_et al., 2011 Figure 1



mmi_7751_f1.pdf

Sharma_et al., 2011 Figure 2



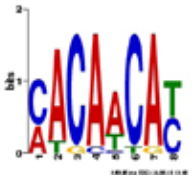
mmi_7751_f2.pdf

Sharma_et al., 2011 Figure 3

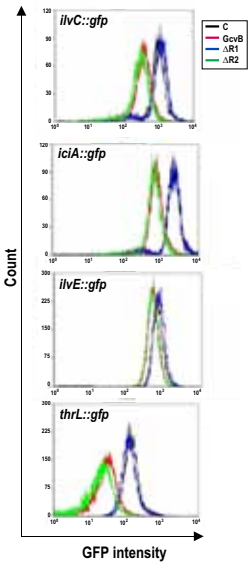
A

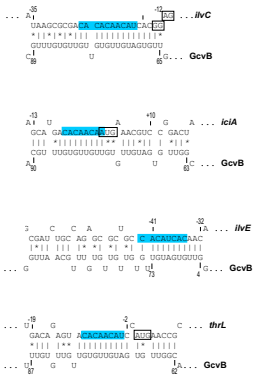
Name	Start	p-value	Sites	
yaeC	10	3.00e-05	TAACGTTAAA	**CACAACAC** AAATACTCAT
livJ	63	3.00e-05	ACAATAACAC	**CACAACAC** ACGTAACAAC
serA	14	6.14e-05	CAAATCCACA	**CACAACAT** CAGATGGCAA
ybdH	65	6.14e-05	CTAAGGACAA	**CACAACAT** GAACCACACT
STM4351	43	6.14e-05	GCACGGTGAC	**CACAACAT** CCAATTGATA
gltI	30	2.03e-04	TGCAGGCAAT	**AACAACAT** CACAATAGCT
dppA	133	2.03e-04	CAAAGGGCAA	**AACAACAT** ACATCACAAT
brnQ	144	2.73e-04	TAATGTTTAC	**AACATCAC** CATCCACAGG
livK	167	2.73e-04	CAAAACATAT	**AACATCAC** GAATGGGGAT
oppA	168	2.73e-04	CACAATGTCT	**AACATCAC** GAAAAAAAGT
tppB	42	3.35e-04	ATGAGTCTGA	**CTCATCAC** GCGCCAGACA
argT	96	4.01e-04	TCTGCAAACA	**CACAACGC** CACGTAAAAC
ygjU	61	5.88e-04	TGACCAGCAA	**CACAATAC** AAGGAATATG
ndk	12	6.77e-04	GACATAACAA	**CAGAACAT** ATTTCAGAGG
gdhA	31	9.42e-04	ACGCAAATAC	**AACCACAT** TAATATATAA
lrp	93	1.00e-03	CAACGGTCTT	**CTCACCAT** AGACCAGGCA

B



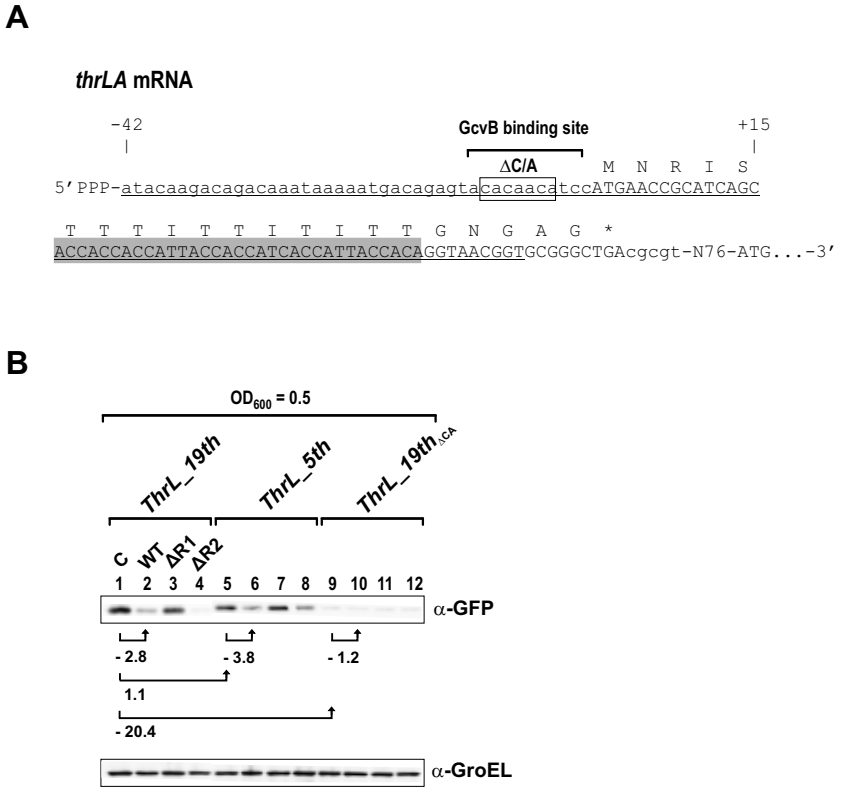
C





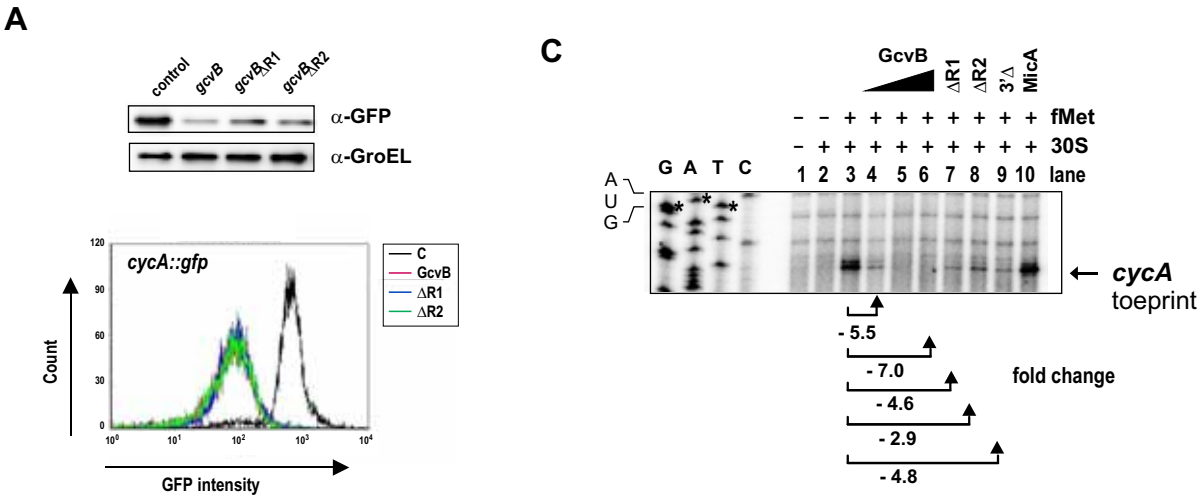
mmi_7751_f3.pdf

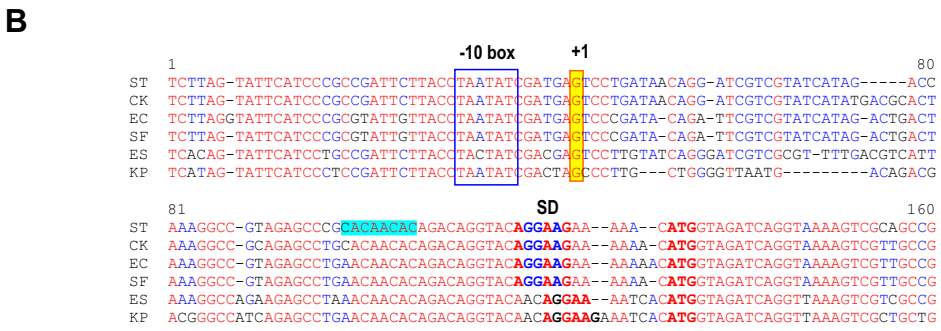
Sharma_et al., 2011 Figure 4



mmi_7751_f4.pdf

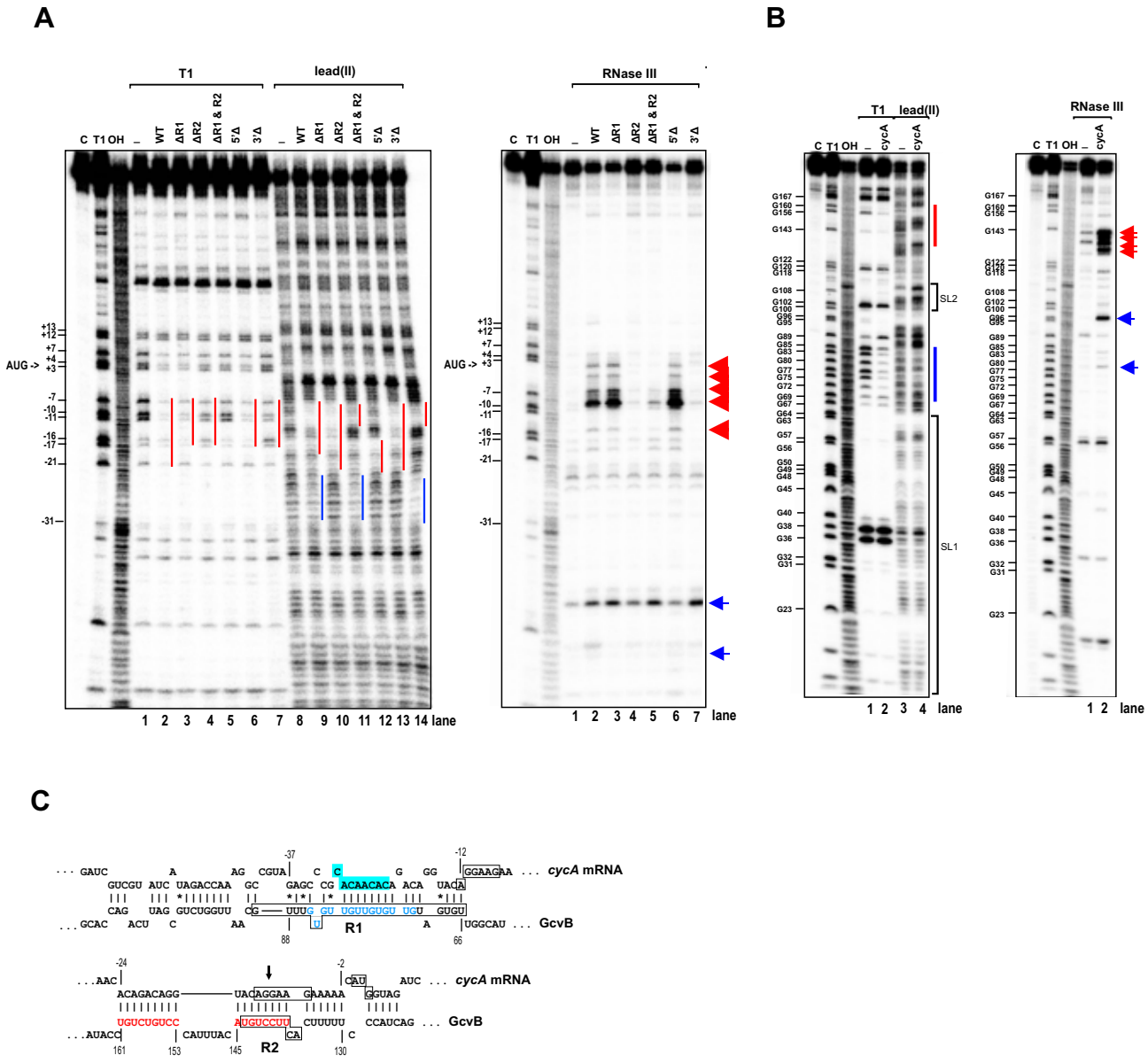
Sharma_et al., 2011 Figure 5





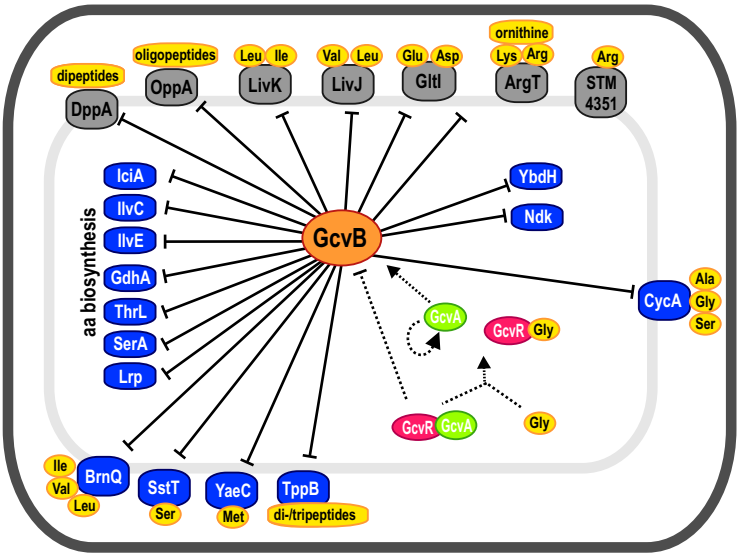
mmi_7751_f5.pdf

Sharma_et al., 2011 Figure 6



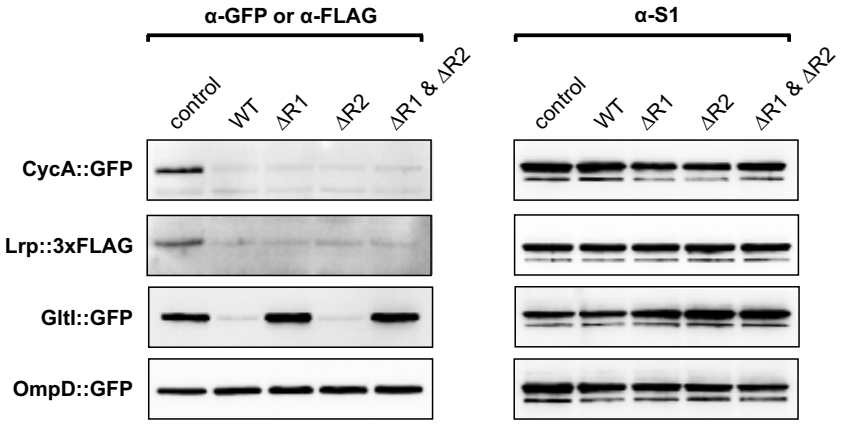
mmi_7751_f6.pdf

Sharma_et al., 2011 Figure 7



mmi_7751_f7.pdf

Sharma_et al., 2011 Figure 8



mmi_7751_f8.pdf